Development of a Self-reported Physical Fitness Questionnaire

CHIK, Tsz Kwan

A Thesis Submitted in Partial Fulfillment
of the Requirement for the Degree of
Master of Philosophy
in
Sports Science and Physical Education

Thesis Assessment Committee

Professor WONG Heung Sang Stephen (Chair)

Professor HUI Sai Chuen (Thesis Supervisor)

Professor Christopher Sean LONSDALE (Committee Member)

Professor ZHANG Jianhui (External Examiner)

Abstract of the thesis entitled:

Development of a Self-Reported Physical Fitness Questionnaire

Submitted by Chik Tsz Kwan

For the degree of Master of Philosophy in Sports Science and Physical Education

At the Chinese University of Hong Kong in August 2007

ABSTRACT

Health-related physical fitness consists of five components and was found to be strongly related to health improvement and disease prevention. Existing methods to assess an individual's physical fitness include laboratory and field tests. Limited studies measuring physical fitness through a self-reported questionnaire were found. The purpose of the study was to develop a self-reported questionnaire for measuring health-related physical fitness. The first stage of the study was to construct as many question items as possible for each component of health-related physical fitness. The initial Self-reported Physical Fitness Questionnaire (SPFQ) was then reviewed by five professionals in the field of physical fitness to examine the content validity (84 items). Criterion-related validity and reliability (internal consistency and test-retest reliability) were evaluated using a sample of 169 Chinese adults who took part in a series of physical fitness tests in the laboratory once and the SPFQ (77 items) twice. Results showed very high test-retest reliability ($r > .84$) and high internal consistency ($\alpha > .85$). Poor criterion-related validity of cardiorespiratory question items was found, but it was fair to moderate for items of muscular endurance, muscular strength, body composition, and flexibility. The criterion-related validity was highest in the body composition component in both trials [women: $r = (-.53) – (-.50)$; men: $r = (-.44) – (-.42)$]. The questionnaire was then revised by removing items that had low correlations with the criterion from each component to improve the criterion-related validity of each component. As a result, the SPFQ was further revised to include only 15 items for women and 9 items for men in the revised SPFQ with four components. The validity coefficients, however, still have room for improvement. Therefore, it can be concluded that future studies are needed to develop a self-reported fitness questionnaire that will include all five health-related physical fitness components, with satisfactory construct validity and norms for score interpretation.

摘要

健康體適能包含五個元素，而且都被發現與改善健康狀況和預防疾病有關。目前用來評估健康體能的方法分爲「實驗室測試」和「實地測試」兩種，而有關於以自我評估問卷形式來評估健康體適能的研究卻非常有限。本研究旨在設計一份能有效地評估健康體適能的自我評估問卷。研究的第一步是爲健康體適能的每一元素盡量構思評估問題（共 84 題）。然後，這份「體適能自我評估問卷」（SPFQ）初稿經五位體適能研究專家評審內容效度，並進行修改，成爲是次研究使用的版本（77 題）。參與是次研究的 169 名香港成年人，需要分別回答這份經內容效度評審後的 SPFQ 兩次，以及參與一系列實驗室體適能測試，以評估其效標關聯效度和信度（內在一致性和再測信度）。分析結果發現 SPFQ 的再測信度很高（$r > .84$），而內在一致性亦然（$\alpha > .85$）。然而與心肺功能有關的題目的效標關聯效度並不理想。至於與肌耐力、肌力、身體脂肪比例和柔軟度有關的題目的效標關聯效度則介乎尚可至中等，當中以身體脂肪比例的題目的效度最高 [女性：$r =$（-.53）–（-.50）；男性：$r =$（-.44）–（-.42）]。及後，研究者對問卷作出修訂，移除了效標關聯效度較低的題目，目的是嘗試改善 SPFQ 中每個體適能元素的效標關聯信度。修訂後，SPFQ 的題目總數減爲女士版本的 15 題和男士版本的 9 題，並且只包含四個健康體適能元素。研究者認爲效標關聯效度仍有改善空間，因此認爲往後除了需發展一份既包含五個健康體適能元素又有建構效度的自我評估問卷，還需訂立問卷得分的規範。

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Background of study

In an article in the magazine Sports Illustrated, John F. Kennedy, the thirty-fourth president of the United States said: "…our increasing lack of physical fitness, is a menace to our national security—such softness on the part of the individual citizen can help to strip and destroy the vitality of a nation" (Kennedy, 1960). President Kennedy foresaw it in the 1960s that physical fitness will be threatened as the "myriad conveniences and distractions of modern life all lure our young people away from the strenuous physical activity that is the basis of fitness in youth and in later life". Although President Kennedy was mainly commenting on the condition of the U.S. Army, the statement is now true for the general public because of its significance to a nation lacking in physical fitness.

Studies have shown a negative correlation between health-related physical fitness and mortality (Blair et al., 1989; Blair et al., 1995; Blair et al., 1996; Ekelund et al., 1988; Erikssen, 1986; FitzGerald et al., 2004; S. Jackson et al., 1995; Katzmarzyk et al., 2004; Slattery & Jacobs, 1988). Being an issue of concern, numerous researchers have strived to develop different ways to measure or evaluate physical fitness. However for the time being, health-related physical fitness components are assessed by time-consuming and costly laboratory and field tests. Although non-exercise models for maximal oxygen uptake (VO_2max) prediction have been established, information such as measurement of body composition and physical activity habits are required, and only the component of aerobic fitness is evaluated. Currently, there are only limited studies that evaluate all components of health-related physical fitness without the need of exercise tests.

Intended for assessing a wider scope of physical fitness, several self-reported questionnaires have been developed (Abadie, 1988; Jackson et al., 1990; Marsh & Redmayne, 1994). In Abadie's questionnaire, four factors of physical fitness, namely, physical condition, muscular strength, muscular flexibility, and body composition, were defined. The 12 questions are rated on a five-point Likert Scale. The rating is based on the individual's subjective perception of one's self rather than the functional ability to carry out daily activities which they would encounter every day. The test-retest reliability was reported to be very high ($r = .92$).

Another similar questionnaire was developed for junior high school girls by Marsh and Redmayne in 1994. Confirmatory factor analysis (CFA) was carried out to classify the six physical self-concept scales and provide support for a multidimensional hierarchical model of physical self-concept. The self-concept was self-reported with a six-point scaled questionnaire, which was correlated with the results from physical fitness tests. However, only three items of self-concept correlated well to physical fitness.

The Physical Activity Rating Scale (PA-R), which is a one-item scale, is indeed one of the variables in an equation for a non-exercise model for aerobic fitness prediction (Jackson et al., 1990). Each participant has to select one number that best represents his/her own general physical activity level for the previous month.

These questionnaires, however, failed to become widely used instruments. One of the limitations is the questions that are asked about an individual's perception on his/her physical fitness instead of the actual functional ability he/she may face in daily living (Abadie, 1988). Furthermore, Marsh and Redmayne's questionnaire only queries the self-concept of adolescent girls' physical fitness levels. This self-concept questionnaire does not necessarily concern activities of daily living (ADLs).

Moreover, the one-item questionnaire developed by Jackson et al. (1990) is only restricted to cardiorespiratory fitness.

In view of the lack of a convenient instrument which allows people to self-assess their physical fitness, the current study proposes to develop a self-reported questionnaire for health-related physical fitness.

Purpose and Significance

This study aims to develop and evaluate the validity and reliability of the Self-reported Physical Fitness Questionnaire (SPFQ) for assessing the health-related physical fitness of Chinese adults.

If the SPFQ is found valid and reliable for the specified age group, it would be a simple, time-efficient, easy-to-administer, self-reliant, economical, and risk-free instrument. Hence, it would be convenient for mass testing and would be useful in epidemiological studies.

Delimitation

Only Chinese adults living in Hong Kong aged 20 to 60 years, both men and women, were included in the present study. Only the five major components of health-related fitness, namely, cardiorespiratory fitness, body composition, muscular strength, muscular endurance, and flexibility, were included in this study.

Limitations

It was assumed that all participants will answer the questions truly and honestly. They were also assumed to have a good understanding of their own physical condition.

Operational Definitions

Health-related physical fitness

Health-related physical fitness refers to the five components of physical conditions that are related to health, as suggested by the American College of Sports Medicine [ACSM] (2000). These components include cardiorespiratory fitness, body composition, muscular strength, muscular endurance, and flexibility.

Cardiorespiratory fitness

Cardiorespiratory fitness refers to the exercise that depends on the functional state of the respiratory, cardiovascular, and skeletal muscle systems.

Body composition

Body composition refers to the extent of body fat relative to fat-free tissue in body weight.

Muscular strength

Muscular strength refers to the maximal force that can be generated by a specific muscle or muscle group.

Muscular endurance

Muscular endurance refers to the ability of a muscle group to execute repeated contractions over a period of time sufficient to cause muscular fatigue, or to maintain a specific percentage of the maximum voluntary contraction for a prolonged period of time.

Flexibility

Flexibility is the ability to move a joint through its complete range of motion.

Self-reported questionnaire

A questionnaire designed for the purpose of being managed and administered by the participants themselves. It is expected that the participants will gain some knowledge from the results of the completed self-reported questionnaire.

Hypotheses

It was hypothesized that the SPFQ would effectively measure the level of health-related physical fitness of Hong Kong Chinese adults with satisfactory reliability and validity.

CHAPTER II

REVIEW OF LITERATURE

A well-designed physical fitness test battery would allow people to learn about their physical fitness level and potential risks for some health problems, as well as to identify their needs. Most of the existing test batteries attempt to measure all or most of the health-related fitness components, but they all require individuals to take part physically in either laboratory or field tests. A self-reported non-exercise fitness assessment battery would minimize the cost and inconvenience associated with laboratory and field fitness tests. However, little research has been conducted in developing such a test. The present study attempts to develop a self-reported non-exercise fitness assessment inventory that can assess the five major components of health-related fitness. To better understand the assessment of health-related fitness, the literature review is divided into six major sections. The evolution of the definition of physical fitness and its association with health and mortality is discussed in the first section, which is then followed by the review on existing health and physical fitness evaluation and programs in different countries and regions. The third section focuses on the existing laboratory and field tests for health-related physical fitness assessment. In the fourth section, several self-reported questionnaire for evaluating physical fitness is reviewed. The fifth section reviews studies concerning the advantages and disadvantages of self-reported questionnaires. The end of this chapter compiles a summary of the previously reviewed studies.

A substantial level of physical fitness brings numerous benefits to an individuals' health (United States Department of Health and Human Services [USDHHS], 1996). In fact, it was not until the 1970s that physical fitness became an issue of concern. The vastly increasing amount of studies on these issues (e.g., Blair, 1993; Blair et al., 2001) revealed that there has been increasing importance and awareness, and thus related scientific investigation, of physical health and fitness in recent decades. Despite the fact that physical fitness is a result of maintaining a certain amount of physical activity, it was found that almost 80% of American adults were not active enough to achieve health benefits (USDHHS, 1996). The finding is more or less the same in Hong Kong where 76% of the Hong Kong people are not active enough to maintain a better health (Hui & Morrow, 2001). Therefore, it is expected that by promoting physical activity, the level of physical fitness would be improved. The public would be more encouraged and convinced if they know their physical fitness level before commencing any physical activity. In other words, the sooner an easy-to-administer physical fitness assessment is developed, the sooner they would know how to improve their fitness level, and the better the public health would be. Consequently, it is obvious that a self-manageable test battery for overall health-related fitness would be needed to monitor people's physical fitness in a large scale.

In fact, there are some national, regional, or organizational batteries that have been developed for the purpose of measuring fitness. Some of these include: the EUROFIT in Europe (Adam et al., 1988; Council of Europe, n.d.); the American Association for Health, Physical Education, and Recreation Youth Fitness Test (AAHPER, 1965), the National Children and Youth Fitness Study (NCYFS) (Federal Prevention Activities, 1994; Pate et al., 1985), and the FITNESSGRAM in the United States (Cooper Institute, n.d.; FITNESSGRAM, n.d.); the National Fitness Assessment in China (China National Sports Council [國家體育總局科教司], 2000); the Physical Fitness Passport in Taiwan (Taiwan Ministry of Education, n.d.); the Canadian Standardized Test of Fitness in Canada (Canadian Fitness and Lifestyle

Research Institute [CFLRI], n.d.); the International Physical Fitness Test in the Middle East regions (Rosandich, 1999, 2003); the Japan Fitness Test in Japan (Japan Ministry of Education, Culture, Sports, Science and Technology [MEXT], 2005), and the YMCA Physical Fitness Test Battery (Golding et al., 1989). Some of these programs are carried out as an award program, such as the President's Challenge Fitness Award Scheme in the United States (The President's Council on Physical Fitness and Sports [PCPFS], n.d.), and many of them were designed for a particular age group or population.

Physical Fitness: Definition and its Association with Health and Mortality

As the concept of physical fitness has become an issue of concern among government and researchers from different countries and regions, it is crucial to have a consistent definition of physical fitness. Physical activity and physical fitness are sometimes used interchangeably, but they actually represent two different ideas. According to the ACSM (2000), physical activity is defined as "bodily movement that is produced by the contraction of skeletal muscle and that substantially increases energy expenditure" (p.4), while physical fitness is "a set of attributes that people have or achieve that relates to the ability to perform physical activity" (p.4). Researchers have been attempting to find out whether activity or fitness is more important for health, but no conclusion has yet been drawn (Blair et al., 2001). These two things together contribute to the prevention of mortality and some chronic diseases.

Although there is a rising concern for health-related physical fitness, the definition of physical fitness has been evolving over the years. Andersen et al. (1996) defined it as the ability to perform muscular work satisfactorily. Marsh and Redmayne (1994) defined physical fitness as a collective term for endurance, balance, flexibility, static strength, and power. Meanwhile, Heyward (2002) explained that it is "the ability to perform occupational, recreational, and daily activities without undue fatigue, which comprises cardiorespiratory endurance, musculoskeletal fitness, body weight and body composition, flexibility, and neuromuscular relaxation". On the other hand, Brock and Fittipaldi-Wert (2005) defined physical fitness as the benefits acquired from the performance. Physical fitness can be further subdivided into motor-related and health-related fitness (Nieman, 1998). When this is the case, balance and power in the definition by Marsh and Redmayne (1994) are just motor-related physical fitness. A few other researchers have also mixed health-related fitness with motor-related fitness. As the study would be lacking a research direction if the researchers did not clearly specify a consistent definition of physical fitness, a standardized and widely-accepted definition is clearly needed.

In 2000, the ACSM standardized the definition of health-related physical fitness as "a state characterized by an ability to perform daily activities with vigor, and a demonstration of traits and capacities that are associated with low risk of premature development of the hypokinetic diseases"(p.57), including components of "cardiorespiratory endurance, body composition, muscular strength and endurance, and flexibility" (p.57). Moreover, to minimize the confusion and ambiguity over them, the definitions for each of the five physical fitness components are clearly stated. Cardiorespiratory fitness is related to the ability to perform large muscle, dynamic, moderate-to-high intensity exercise for prolonged periods … the performance of such exercise depends on the functional state of the respiratory, cardiovascular, and skeletal muscle systems (p.68). Muscular strength refers to the maximal force that can be generated by a specific muscle or muscle group (p. 81). Muscular endurance is the ability of a muscle group to execute repeated contractions over a period of time sufficient to cause muscular fatigue, or to maintain a position for a specific period of time (p. 84). Flexibility is the ability to move a joint through its complete range of motion (p. 85). Body composition refers to the relative percentage of body weight that is fat and fat-free tissue (p. 60).

ACSM's definition of physical fitness and its components are more inclusive and widely recognized than others. The following review of literature will be based on the five components of physical fitness, as defined by the ACSM: cardiorespiratory, muscular endurance, muscular strength, flexibility, and body composition.

Brief Review on Physical Fitness and Health

A strong association was reported between physical fitness maintenance and improvement and all-cause mortality in American men and women (Blair et al., 1996; Blair et al., 1989; Blair et al., 1995; Blair et al., 1998; DiPietro et al., 1998; Katzmarzyk et al., 2004; Villeneuve et al., 1998). Blair showed in his study that physical fitness has a stronger inverse association with mortality than physical activity (Blair, 1993). Substantial cardiorespiratory fitness could protect healthy men, as well as men with the metabolic syndrome, from all-cause and cardiovascular disease mortality (Katzmarzyk et al., 2004). As a strong predictor of mortality, low fitness is as risky as smoking (Blair et al., 1998). Blair and his colleagues compared persistently unfit men with men who had their physical fitness improved (Blair et al., 1995). Results revealed that men who had an improvement in physical fitness had a reduction of 44% in premature mortality risk. More importantly, men who have been maintaining sufficient physical fitness had the lowest death rate. Blair and his colleagues conducted an observation of 25,341 American men and 7,080 American women and realized that fit people with any premature mortality risk factors (including smoking, elevated blood pressure, and cholesterol level) had lower death rates than low-fit people without any risk factors (Blair et al., 1996).

The association of cardiorespiratory fitness with coronary heart disease is obviously inversed, regardless of race or occupation (USDHHS, 1996). It was also found that both fitness and physical activity give further protection against coronary heart disease by modifying risk factors, such as serum insulin, obesities, and serum sex hormones (Tikkanen et al., 1998). The study indicated that the better an individual's cardiorespiratory fitness is, the less likely the risk of coronary heart disease risk factors. This is true even in a low-risk sample of the middle-aged individuals. However, no studies have been found that studied the association between health-related fitness and strokes.

Other than cardiovascular fitness, muscular fitness in terms of muscular strength and endurance is also found to be essential in increasing the ability to perform ADLs and overall quality of life, and in decreasing the prevalence of disability, morbidity, and premature mortality, especially among women (Warburton et al., 2001). Having good muscular fitness may also reduce incidences of falling and its associated injuries. The researchers also mentioned that substantial flexibility could increase mobility and functional independence, which is particularly true in the elderly. Regarding musculoskeletal function, studies show that enhanced muscle fitness may strengthen the neuromuscular and skeletal systems and decrease the risk of injury (Hootman et al., 2001; Tikkanen et al., 1998). A common problem amongst the middle-aged is osteoporosis. It is known that insufficient weight-bearing physical activities are one of the factors that cause osteoporosis in females in their 70's. However, a British study gave a new perspective by suggesting that maintaining high levels of cardiovascular fitness by performing various aerobic activities would contribute to good bone health (Turner, 2000).

Few studies have examined the association between physical fitness and mental health, although it is reported that physical activity improves physical fitness and mental health and may thus protect against the development of depression (USDHHS, 1996). The U.S. Surgeon General's Report indicated that depression and anxiety could be relieved by regular physical activity. Exercising also provides a release from tensions in life (Wuest & Bucher, 2003). Either aerobic (eight-week running) or anaerobic (weight-lifting) exercises could significantly reduce depression in women who were diagnosed with depression in the study (Doyne et al., 1987).

Physical Fitness Evaluation Programs

Because of the importance of physical fitness, the implementation of corresponding evaluation is urgently needed. If people are informed of their level of physical fitness and health risk, corresponding recommendations for improvement can be prescribed (Stuck et al., 2002). For this reason, a number of health and physical fitness evaluation programs have been developed in different countries and regions.

Since the early sixties, several well-known fitness evaluation programs in the United States have been developed: AAHPER Youth Fitness Test, National Children and Youth Fitness Study (NCYFS), FITNESSGRAM (Cooper Institute, n.d.; FITNESSGRAM, n.d.), and The President's Challenge. AAHPER Youth Fitness Test is the first comprehensive test battery ever developed (AAHPER, 1965). Initiated in 1957, the manual with national norms was finalized in 1965. The AAHPER developed this battery to measure the physical fitness level of the fifth to twelfth grade American youth. Seven items are included in the test battery: pull-up (boys) or flexed-arm hang (girls), 1-min sit-up, shuttle run, standing long jump, 50-yard dash, softball throw for distance, and the 600-yard run-walk. These items do not only measure health-related fitness, but also motor fitness. The battery has gained popularity among American physical education teachers and many youth-serving organizations throughout the United States and some foreign countries since it was developed.

First published in 1985 (Pate et al., 1985), the National Children and Youth Fitness Study (NCYFS I) was soon modified as the NCYFS II in 1987 (Federal Prevention Activities, 1994). This nationwide study was implemented under the auspices of the U.S. Public Health Service. Children aged 10-18 years old and 6-9 years old drawn from 25 counties in the United States were tested on nine fitness measures. Norms were generated (Safrit & Wood, 1995).

In 1982, the Cooper Institute in Dallas, Texas, developed the FITNESSGRAM (Cooper Institute, n.d.; FITNESSGRAM, n.d.). It is the first complete assessment of health-related physical fitness and activity for children and youth that uses a criterion-referenced standard. It aims to increase parent's awareness of the fitness level of their children. The program was designed to be easily manageable so that school teachers could use them. Various options are available for assessing all five physical fitness components as suggested by the ACSM. Aerobic capacity is measured by a PACER test (recommended), 1-mile run/walk, or walk test for ages 13 or above. Body composition is calculated with two-site skinfolds (recommended) or body mass index (BMI). Muscular strength, endurance, and flexibility are measured with curl-ups, trunk lifts; either one from a push-up, a modified pull-up, and a flexed arm hang; and either one from a back-saver sit-and-reach and a shoulder stretch. The first three items were recommended for assessing muscle fitness and flexibility. Results would be compared to the standards set for genders and ages. This computer program might be an innovation to produce an individualized easy-to-understand report after the test for each youth participant and his/her parents that explains their fitness condition and gives suggestions for health improvement. This program is applicable to a large population and is suitable even for young adults (age $\leq$ 30 years old). The interface design and capability of the program is still being improved with different versions on the market (Mahar et al., 1997).

The President's Challenge, started with the Presidential Physical Fitness Award in 1966, was set up to motivate American youth (age = 6-17) to exercise for fitness and to assess their current fitness level (PCPFS, n.d.). Three program areas are included in the President's Challenge: Active Lifestyle, Physical Fitness, and Health Fitness. The program of Physical Fitness can be further split into three levels according to the performance level in the physical fitness program, namely, Presidential Physical Fitness Award (PPFA), National Physical Fitness Award (NPFA), and Participant Award (PA). In recognition of the participants whose fitness level is regarded as healthy in all events of this Program, the Health Fitness Award (HFA) also evaluates using the BMI. The Physical Fitness Program assesses abdominal strength and endurance (curl-ups or partial curl-ups), total body coordination (shuttle run), cardiorespiratory endurance (one-mile run; or 0.25 mile for 6-7 years old; or 0.5 mile for 8-9 years old), upper body strength and endurance (pull-ups or right angle push-ups), and muscular flexibility (V-sit reach or sit-and-reach). This award program does not only successfully promote physical fitness to the youth, but can also be applied to all ages of the population, from 6 to older than 65 years old.

The development of EUROFIT was began in 1978 by the Committee for the Development of Sport in Europe to test the physical fitness of European school-age children and adolescents, aged between 6 and 18 (Adam et al., 1988; Council of Europe, n. d.). It was not until 1986 when the final approval was made by the Committee of Ministers. EUROFIT was based on the aims of setting up a commonly agreed test battery in Europe, helping teachers assess the physical fitness of their students, and helping in measuring the health-related fitness of the population. Testing items include: height, weight, skinfolds measurement from four sites, and eight items concerning "motor fitness". These eight items include the Flamingo balance test (balance), plate tapping (limb agility), sit and reach (flexibility), standing broad jump (explosive strength), hand grip (static muscular strength), sit-ups (abdominal muscular strength), bent arm hang (arm and shoulder muscular endurance), and 10 x 5M-shuttle run (agility). To measure cardiorespiratory endurance, two indoor tests, namely, endurance shuttle run and bicycle ergometer test, were also available for use with the above test items.

The United States Sports Academy developed the International Physical Fitness Test for Arabian youth aged 9 to 19 (Rosandich, 1999, 2003). The project was in collaboration with the Supreme Council for Youth and Sport, now known as the General Organization of Youth and Sport. Initially developed in 1964, the results were first presented in 1977. This test battery had five components: 50-M dash test (relative power and speed), pull-up (strength), 10-M shuttle (relative power, speed, and agility), back throw (power, speed, and agility), and 1000-M run (aerobic/ anaerobic capacity and stamina). The test, however, measures with similar running elements repeatedly merely for speed, strength, suppleness (agility), and stamina (endurance). Although it had been modified to a one-day test, running more than one kilometer could be harsh work for the youths who are physically unfit, especially under the hot weather of the Middle East.

Developed in 1989, the YMCA Physical Fitness Test Battery measures all of the five health-related physical fitness components suggested by the ACSM. It was intended to measure the physical fitness of the YMCA adult members. Bike submaximal test or three-minute step test (cardiorespiratory capacity), three- (female) or four-site (male) skinfold test (body composition), sit-and-reach test (flexibility), and bench press test or one-minute half sit-up test (muscular strength and endurance) were used as test items. Age- and gender-specific norms were developed for assessing individual fitness level (Golding, 2000; Morrow et al., 2000). These test items are still widely used.

Eight years before the YMCA Physical Fitness Test Battery was developed, the first Canadian Standardized Test of Fitness was developed by the Canadian Fitness and Lifestyle Research Institute (CFLRI) as part of the 1981 Canada Fitness Survey by the Canadian Society for Exercise Physiology [CSEP] (CFLRI, n.d.). It was designed for all Canadians aged 7 and above living in households. Physical fitness in terms of cardiovascular fitness, flexibility, muscular endurance, and strength, was evaluated. The most recent edition was published in 1998: the Canadian Physical Activity, Fitness and Lifestyle Appraisal (CPAFLA) by the CSEP (1998).

Japan also has a similar program called 体力・運動能力調査, or the Tairyoku Undou Nouryoku Chousa (Japan Fitness Test) (MEXT, 2005), as part of the Physical Strength and Motor Fitness Survey. It began in 1998 to assess the level of physical fitness among citizens aged 6 to 79 by the MEXT. Four sets of fitness tests were created for primary school pupils (6-11 years old), secondary to tertiary school students (12-19 years old), adults (20-64 years old), and elderly (65-79 years old), respectively. The age group division of adults in this Japanese test battery is different from other countries due to their cultural tradition that classifies adults as

20-year-olds and above. Handgrips, sit-ups, and sit-and-reach were used in all age groups. Single-leg balance (with eyes closed), 6-minute walk, 10-m walk with obstacles, and a questionnaire of ADLs were only given to the elderly group. On top of these required items, the youngest group (age = 6-11) also had to perform a side step test, 20-m shuttle run, 50-m run, standing long jump, and softball throw. Then, for the group of secondary school and college or university students, the softball throw was replaced by the handball throw, and an item of endurance run was added. The adult group had the least number of items: 20-m shuttle or rapid run, plus the standing long jump. The annual assessment has been in operation until now.

The Taiwan Physical Fitness Passport was initiated by the Taiwan Ministry of Education (n.d.). This Passport, similar to the FITNESSGRAM, aims to assess the health-related physical fitness of children and youth. It is also similar to the President's Challenge in the way that students will be awarded by attaining a certain level of physical fitness. In 1999, the scheme was experimented given to secondary and primary school students, and in the following two years, to college/ university students and teachers (Jwo, 2000).

To assess and improve the physical fitness of the Chinese people in Mainland China, the Chinese government implemented the first physical fitness assessment in 2000 to Chinese people of all ages (China National Sports Council [國家體育總局科教司], 2000). Four groups were divided according to their age: preschool children (0-6 years old), children and youth (7-18 years old), adults (male: 19-59 years old; female: 19-54 years old), and elderly (male: 60 years old or above; female: 55 years old or above). All five fitness components were tested with different test items for different age groups. A norm table was created after the assessment for further physical fitness evaluation.

Existing Laboratory and Field Tests

Physical fitness could be assessed by a laboratory test, a field test, or a non-exercise prediction test. These methods vary in required time, cost, skill, and, hence, applicability, accuracy, and universality. The following section examines some of the more important test protocols.

Laboratory Methods

Laboratory tests are available for assessing the five physical fitness components. They usually require trained personnel (e.g., trained tester, clinical supervisor), sophisticated and expensive equipment and set-up, and more time, yet potential risks cannot be avoided. To minimize risks, some safety measures and guidelines should be given, and the individual's health and medical information should be obtained before physical fitness testing. Apart from screening the participants and distributing an informed consent and the PAR-Q, individuals should also be given precise instructions before the fitness testing, such as suitable clothing and what to and what not to eat or drink before the testing. Should these safety measures be carefully observed, laboratory tests can be safely carried out (ACSM, 2000).

Cardiorespiratory fitness. The standard measure of cardiorespiratory fitness is the maximal oxygen consumption (VO_2max). It is the product of maximal cardiac output and arterial-venous oxygen difference, that is, VO_2max related to the functional capacity of the heart. It requires the participants to exercise to exhaustion by regularly increasing the power output. A submaximal test, on the other hand, assumes a steady heart rate to be maintained during each exercise test, a positive relationship between heart rate and work rate, a uniformed maximal heart rate according to age, and the VO_2 at a given work rate is the same for all people. The judgment on whether to use a maximal or a submaximal test relies on the type of participant, the reasons for the test, and the availability of appropriate equipment and personnel (ACSM, 2000).

Although VO_2max can be directly measured in the laboratory by collecting the expired gas using the Douglas bags and chemical gas analyzers to determine the oxygen and carbon dioxide content, sophisticated laboratory equipment, skills, and time are required (Williford et al., 1996). All these have prohibited it from gaining further popularity. VO_2max is usually measured on an ergometer, which includes treadmill running or walking and bicycle. Several maximal and submaximal test protocols are available for treadmills and bikes, such as the Åstrand-Ryhming single-stage test which lasts for six minutes. To know whether an individual's heart rate lies within the target heart rate of 125-170 beats per minute (bpm), the obtained post-exercise heart rate can be referred to the age-adjusted Åstrand-Ryhming nomogram (Åstrand & Ryhming, 1954).

Body composition. Major laboratory methods for measuring body composition include hydrostatic weighing and air displacement plethysmography (ADP), which are densitometry methods, as well as dual-energy X-ray absorptiometry (DXA). Hydrostatic weighing is the most popular laboratory method for body composition and is regarded as the golden standard for assessing body density (ACSM, 2000). It determines the body volume according to Archimedes' principle. By knowing the body weight and volume, body density can be obtained by estimating the percent body fat (%BF) from Siri's equation:

$$\%BF = (495/BD) - 450$$

Several race-specific formulas are also available for the conversion of body density to estimate %BF (Heyward, 2002).

ADP works principally similar to hydrostatic weighing, except that air displacement method uses air that the human body has displaced when it enters the container. After the body volume and the body density are determined, body composition can be calculated. Validity and reliability of ADP, such as the Bod Pod, was found satisfactory to assess body composition in various populations (Ball, 2005; Biaggi (Ball, 2005; Biaggi et al., 1999; Bosy-Westphal et al., 2004; Clasey & Gater, 2005; Demerath et al., 2002; Flakoll et al., 2004; Frisard et al., 2005; Koda et al., 2000; Levenhagen et al., 1999; Maddalozzo et al., 2002; McCrory et al., 1995). Some studies even found it more valid and reliable than other methods (Demerath et al., 2002; Sardinha et al., 1998; Vescovi et al., 2001). ADP can obtain the body composition fast and is relatively easy-to-operate with reasonable cost (Flakoll et al., 2004; McCrory et al., 1995; Radley et al., 2003). Its non-invasive procedure could accommodate even special populations such as the elderly and the disabled (Bosy-Westphal et al., 2004).

DXA measures body composition including body fat, bone mineral, and lean mass. Satisfactory results of correlation to hydrostatic weighing and validation of detecting small changes in body composition were reported in previous studies (Going et al., 1993; VanLoan & Mayclin, 1992). DXA gained agreement from some studies that revealed reasonable accuracy and reliability (Flakoll et al., 2004; Glickman et al., 2004; Maddalozzo et al., 2002; St-Onge et al., 2004).

Muscular fitness. Muscular fitness includes muscular strength and endurance, but none of them can be tested with a single test (ACSM, 2000). In a laboratory, a popular assessment for muscular strength and endurance is isokinetic testing, in which a constant angular velocity is set for maximal muscular force throughout the limb's range of motion (Heyward, 2002). Because a computerized isokinetic dynamometer is the most valid method, it is often used for academic research and clinical purposes.

Flexibility. Flexibility is specific to the joints involved. Common laboratory methods include direct measurements using goniometers, flexometer, and inclinometers and indirect measurements using visual estimation. They can measure the angle of the range of motion of a specific joint one by one (ACSM, 2000).

Field Tests

Field tests usually require less equipment and take less time than laboratory tests. They are therefore more commonly used for fitness assessment than laboratory tests.

Cardiorespiratory capacity. Five commonly used tests for VO_2max prediction are: the Cooper 12-minute run test, the Rockport One-Mile Fitness Walking Test, 1.5-mile run test, 1-mile jog test, and 1-mile run/walk test. The Cooper 12-minute run test requires participants to cover as much distance as possible within 12 minutes (Cooper, 1968). The Rockport One-Mile Fitness Walking Test was developed for adults aged 20 to 69 (Kline et al., 1987). It is safe for sedentary adults because it requires only fast walking for one mile. Heart rate (HR) is to be measured in the final minute during the last quarter mile, or a 10-second HR immediately after the 1-mile walk. George and his colleagues had two similar run tests with acceptable accuracy in predicting VO_2max. The 1.5-mile run test requires the participants to run the destined 1.5 mile in the shortest possible time. The 1.5-mile run would be more suitable for individuals who have regular exercises. An alternative would be a submaximal 1-mile track jog (George et al., 1993). Male participants are to finish jogging one mile in not more than eight minutes and female for not more than nine minutes. The HR of all participants should be kept at no greater than 180 bpm. It would be more feasible for all ages and people with different physical abilities.

Another 1-mile run/walk test was for predicting VO_2max in youth and young adults (Cureton et al., 1995). Participants had to run as fast as they could for one mile. A summary of the five field tests are shown in Table 1.

Table 1

Summary of common cardiorespiratory tests

Study	Test	N	Predictor variables [a]	Age group	r
(Cooper, 1968)	12-min run	115	Walk-run distance	U.S. Air Force male officers and airmen (17-52 yr)	.90
(Kline et al., 1987)	1-mile walk	343	Wt, age, G	Healthy adults (20-69 yr)	.92
(George et al., 1993)	1-mile track jog	54	Wt, G, jog time, HR	Relative fit college students (18-29 yr)	.87
(George et al., 1993)	1.5-mile run	50	Wt, G, jog time	Relative fit college students (18-29 yr)	.90
(Cureton et al., 1995)	1.5-mile run/ walk	495	G, BMI, age, MRW	Youth & young adults (8-25 yr)	.72

[a] Predictor variables: Wt, weight (kg); G, gender; jogging time for 1.5 mile (min/ mile); BMI, body mass index; MRW, 1-mile run/walk.

These five tests share several similar attributes. Except for a model by George and his colleagues in 1993, all of them require either a maximal distance within a given time or a minimum time for a given distance. They displayed high correlation ($r \geq .72$) with the criterion measures. These tests do not require a large amount of expensive and bulky equipment, only stopwatches and the ground for walking or running. The most important factor is that a large group of participants can be handled at one time (ACSM, 2000); it is therefore economically efficient.

Body composition. The rationale of the skinfold test is that the total BF can be estimated by measuring the thickness of the subcutaneous fat at several sites. Related studies could be dated back as early as 1967 when the researchers reported a high correlation between skinfold thicknesses and density for both women ($r = -.78$) and men ($r = -.84$) (Durnin & Rahaman, 1967). Specific skinfold sites and procedures are described by the ACSM and various gender-specific generalized skinfold equations are also listed (ACSM, 2000).

Another non-invasive method is the bioelectrical impedance analysis (BIA). It is a device that allows a small electric current to pass through the body and measures the impedance or opposition to electric current flow (ACSM, 2000). The higher content of total body water (TBW) an individual has, as electrolytes in the body's water conduct electricity, the less resistance the current flow has. Thus, the individual would be estimated as leaner. However, it is more accurate than DXA to assess overweight adults (Frisard et al., 2005).

Other anthropometric methods, such as BMI and waist-to-hip ratio (WHR), have also been developed for body composition assessment as field tests. BMI is obtained by having kilograms of body weight divided by squared meters of height (kg/m^2). The index increases with body fat content and obesity-related health problems. However, the traditional western standard may not be applicable to Chinese people. For this reason, the World Health Organization [WHO] reported a revised standard for Asian people. A person is classified as overweight when BMI is 23 kg/m^2, and is obese when BMI is 25 kg/m^2 (WHO, 2000). The WHR, on the other hand, represents the circumference ratio between the waist and the hip, and may therefore be better able to describe the distribution of body fat than the BMI. The ratio is calculated when the circumference of the waist is divided by the circumference of the hips. Similar to the BMI, a higher WHR implies a higher proneness to health problems, but gender has to be considered for both calculations (ACSM, 2000). Interestingly, there is an increasing concern in the use of waist circumference (WC) as a measure of health index. It was found that WC could be a health predictor (Shibuya et al., 2005), even better than merely the BMI (Wildman et al., 2005; Zhu et al., 2005).

Muscular condition. Various testing procedures and equipment are available for assessing isometric and dynamic strength. One common method to assess isometric strength is to use a respective isometric dynamometer that can assess the strength of the leg, back, or grip (Heyward, 2002). Norms of grip strength of each hand are accessible from the Canadian Physical Activity, Fitness and Lifestyle Appraisal [CPAFLA] (CSEP, 1998). Total strength is obtained by summing up right and left grips, leg strength, and back strength scores. Another popular method is a dynamic strength test involving bodily movement. It is usually measured as a 1-repetition maximum (1-RM), such as the bench press (Golding, 2000) and the leg press for upper- and lower-body endurance, respectively, which are norm-referenced (ACSM, 2000). Upper-body strength can also be measured with a consecutive push-up test (use knees as pivot for females) that is used in the Canadian Standardized Test of Fitness (CFLRI, n.d.). All these tests require the participant to perform some warm-up beforehand to avoid injury from a sudden muscle contraction. Although a 1-RM can be estimated from a 6- or 10-RM test, according to the ACSM (2000), the true 1-RM should be measured. Similarly, partial curl-up (CSEP, 1998) and YMCA's

1-minute timed sit-up test (Golding, 2000) measure abdominal endurance. In the former, as many consecutive curl-ups as possible are performed to exhaustion according to a cadence of 25 bpm, while performing the latter is restricted to one minute. Instructions are given in the respective guidebooks.

Flexibility. It is more popular to use the norm-referenced sit-and-reach test than any laboratory method to measure flexibility. Various sit-and-reach test protocols have been developed over the years to measure the low back and hamstring flexibility. The most standard sit-and-reach test is to use a sit-and-reach box. The client has to extend two knees on the floor and lean forward with both hands sliding along the scale on the box (CSEP, 1998). Similarly, YMCA's version is to perform with a yardstick on the floor instead of the box (Golding, 2000). A back-saver sit-and-reach test (Hoeger & Hopkins, 1992) and its modified version (Hui & Yuen, 2000) were devised to minimize the discomfort that occurs during the test, with a sit-and-reach box and on a bench, respectively. These tests measure one leg at a time while the other leg is flexed. A warm-up exercise before these tests is recommended, and a slow and smooth movement could prevent injury (ACSM, 2000).

Problems Associated with Laboratory and Field Tests

Despite the usages and advantages of the laboratory and field tests discussed, most of them carry some disadvantages that discourage some individuals from using them. ACSM (2000) summarized several disadvantages of maximal and submaximal exercise tests for cardiorespiratory fitness assessment. It is said that the maximal test, which requires the participants to exercise to exhaustion, increases the risk of injuries. Physician supervision and emergency equipment may therefore be required during the assessment. Due to its intense exercise, only limited populations are suitable for the assessment. All these mean that a maximal test is not feasible for most health or fitness practitioners. Although submaximal tests provide a safer environment for a participant, precision is sacrificed. The bike tests, either maximal or submaximal, may not be suitable for those unfamiliar to cycling, causing localized muscle fatigue (ACSM, 2000). The test participant may stop the test merely due to muscle fatigue, thus not being able to accurately predict his/her VO_2max. For the field tests, because the 1.5-mile run test is running for time, an individual's performance is highly reliant on motivation and pacing ability. Therefore, it is not conclusive that the 1.5-mile run is the most accurate test for cardiorespiratory fitness estimation.

For body composition, although hydrostatic weighing is regarded as the gold standard for measuring body composition, the use of it is hindered by several factors. Equipment is the most important concern. It requires a pool or a tank for the test, which means the organizers of the test have to be able to afford the corresponding space and piping system. Compared with other methods, a trained operator is also needed to carry out the procedures, and it takes a longer time to perform (Vescovi et al., 2001). This method would be especially problematic for infants, young children, older people, or those who are water-phobic.

Appraisals on DXA were indeed two-sided. Despite the merits discussed earlier, some studies suggested that DXA alone is not as accurate as when other methods are used together for prediction (Johansson et al., 1993; Panotopoulos et al., 2001; Svendsen et al., 1991). Moreover, different DXA systems were believed to give different results from body composition measurement (Norcross & Van Loan, 2004; Soriano et al., 2004; Yang et al., 2004). Because its validity and reliability of precisely estimating percent body fat are still questionable (Bolanowski & Nilsson, 2001; Hansen et al., 1999; Houtkooper et al., 2000; Kitano et al., 2001; Koda et al., 2000; Kohrt, 1995), DXA has yet to be the best referenced method.

Some researchers doubted the accuracy or validity of estimating %BF with ADP (Ball & Altena, 2004; Wagner et al., 2000). Others maintained that it is not consistently accurate in people of different genders, ages, body fat content, and health conditions (Flakoll et al., 2004; Levenhagen et al., 1999; Radley et al., 2003; Vescovi et al., 2001); also suggesting that ADP cannot be widely accepted for all ethnic and age groups.

It has been shown that BIA could not accurately estimate %BF (Flakoll et al., 2004) due to the difference in body water content. In addition, instructions for the participants, such as fasting and avoiding certain food before certain hours of assessment, are hard to ensure (ACSM, 2000). Failure to meet those guidelines would affect the results.

Although BMI is just an objective measure, it does not indicate body fat distribution. The commonly used standard also cannot be applied to all races. As for the Hong Kong Chinese population, it was found that people had a higher %BF for a given BMI (He et al., 2001). The BMI therefore cannot be a universal index. Measurement errors may occur in heavily muscled and obese people as the underlying muscle and fat are squeezed under the skin (Heyward, 2002).

Not many choices of well-established methods are available for muscular fitness assessment. The common methods are isokinetic, isometric, and dynamic strength testing methods. Although a computerized isokinetic dynamometer is considered the most accurate among several methods, the high cost and inconvenience with which it is associated makes isokinetic testing infeasible and impractical for mass testing. Hence, field tests are usually more preferable to others for muscular fitness evaluation.

When measuring flexibility, using apparatus such as goniometers, flexometers, and inclinometers are accurate, but the cost of the equipment decreases people's motivation and demands for utilizing them. Also, specific equipment and trained administrators needed for these direct measures would definitely increase the costs of implementing flexibility assessment. Visual estimation is much less costly but is more likely to provide an inaccurate estimate (ACSM, 2000).

In short, all laboratory and field tests are not convenient for everyone. Demands for equipment and facility, time, and financial support do not make them affordable and accessible to some researchers. In addition, as many testing items in either laboratory or field tests were drawn from some test batteries, test results must be carefully interpreted (Heyward, 2002).

Non-Exercise Prediction Models

In 1990, Jackson and colleagues developed two non-exercise models (N-Ex models) that could estimate VO_2max for adults without involving exercise testing. The N-Ex models only require gender, age, body composition, and self-reported activity as predictor variables. The original 8-point physical activity scale (PA-R) scale (0-7) provided was developed for use at the National Aeronautics and Space Administration Johnson Space Center [NASA/JSC] in Houston, Texas. The participants in this study rated their physical activity level in the previous month based on PA-R by selecting one number that best represents their physical activity level. Two models, representing body composition with %BF from a summation of three-site skinfolds (N-Ex %BF) and BMI (N-Ex BMI), respectively, were developed in this study (Jackson et al., 1990). It was found in validity analyses that the accuracy exceeded the Åstrand submaximal models that systematically underestimated VO_2peak, and was only slightly less accurate than the Rockport walking test. The multiple correlations of N-Ex %BF and N-Ex BMI were .81 and .78, respectively.

Although the N-Ex BMI is slightly less accurate than N-Ex %BF, the models are appropriate for about 96% of the participants in the population studied, and can be generalized to women and men. However, both models were not applicable to the most aerobically fit individuals (VO_2 peak$\geq$ 55 ml $\cdot$ kg^{-1} $\cdot$ min^{-1}).

The N-Ex models have been used to conduct research on a university student population (n = 69) for prediction of aerobic capacity, but the result was not satisfactory (Kolkhorst & Dolgener, 1994). The low predictive value by the N-Ex equations among university students was revealed in high total error values. The divergence between the predicted and measured means of VO_2 peak, according to the investigators, is a possible result of the lack of available validity and reliability information of the PA-R scale, different interpretations on exercise intensity between two subject groups, the close-to-examination time when the study was being carried out, a lower percentage of males, a younger and more homogeneous sample whose aerobic fitness is above average, and a lower self-reported activity level. The authors therefore concluded that the original N-Ex prediction equations underestimated aerobic capacity in aerobically fit university students.

Subsequent to the research described above, George, Stone, and Burkett attempted to predict VO_2max with gender, perceived functional ability (PFA) to walk, jog, or run given distances, modified 10-point Physical Activity Rating Question in N-Ex prediction models (7-point scale) by A. S. Jackson et al., and BMI as variables (George et al., 1997). Despite being unique and having a compatible accuracy to estimate VO_2max entirely from questionnaire-based data, the model in question was only found to be valid and convenient for physically active college students (R = .85, SEE = 3.43-3.47 ml$\cdot kg^{-1} \cdot min^{-1}$), and no significant difference existed between male and female models. Its cross-validity with A.S. Jackson and colleagues' models demonstrated high accuracy (r = .71, SEE = 4.62). However, the tested group was homogeneous in terms of age, physical fitness, and educational level. Thus, according to George and his colleagues, whether or not this questionnaire-based N-EX regression model is valid for other ages is not yet known (1997).

Self-Reported Questionnaire

Self-reported questionnaires may be an alternative to the above methods to assess an individual's physical fitness. Although some similar self-reported questionnaires exist, none has been developed to comprehensively assess all of the five physical fitness components.

Perceived Physical Fitness Scale

Abadie (1988) attempted to develop a more inclusive self-assessed instrument to evaluate the effect of an individual's perception of his/her physical fitness on psychological functioning. Each of the five areas in physical fitness, subsequently adjusted to four, included physical condition, muscular strength, muscular flexibility, and body composition. Each area was addressed by using at least two questions regarding daily functional activities for adults. The questions are rated on a five-point Likert Scale. The results of the corresponding content validity, construct validity, and concurrent validity are discussed in the section of "Validity and Reliability of the Self-Reported Fitness Questionnaires".

NASA/JSC Physical Activity Scale (PA-R)

This scale is not a "stand-alone" survey; on the contrary, the results were used as one of the variables in the gender-specific equations for VO_2max estimation (Jackson et al., 1990). The scale was originally developed for use at NASA/JSC (Wier et al., 1989). The PA-R consists of only one question that aims to survey the physical activity of the employee in NASA. Participants in Jackson et al.'s study were required to select one number that best represents their general physical activity level for the previous month. Descriptions and examples of physical activity were given for each category. Unfortunately, the non-exercise test only predicted aerobic fitness even though high accuracy was established.

Multidimensional Physical Self-Concept

A study conducted in Australia examined the relationship between self-concept of physical fitness and the assessed physical fitness (Marsh & Redmayne, 1994). The self-concept was self-reported with a six-point scaled questionnaire that was distributed to 105 junior high school girls. Each self-concept consists of eight to ten statements. Physical fitness tests include Cooper's 12-minute run test (1968) for endurance, handgrip test for static strength, closed-eye standing on one leg for balance, shuttle run test for strength/power and dynamic flexibility. The participants had to bend forward to touch the floor, straighten up, and twist to the left and the back for as many times as possible in 20 minutes.

Validity and Reliability of the Self-Reported Fitness Questionnaires

In Abadie's study (1988), the relevance of the items was tested by the content validity test. The content validity of the initial 15 items was done by three judges. The judges were to evaluate the items by giving a five-point rating according to the relevance to physical fitness, with "5" being the highest association and "1" being the least association. Twelve items receiving average ratings of four or higher ("considerably related to physical fitness" and "greatly associated with physical fitness") were retained for the scale.

Item analyses showed that the item-to-scale correlation was $r > .20$. Two samples were classified by age: older adults (> 50 yr.) and younger adults (< 50 yr.). Scale factor scores were correlated with muscular strength. While the scale received moderate validity (Table 2) and the test-retest reliability was reported as very high ($r = .92$), the rating is based on the individual's subjective perception on oneself rather than the actual ability to carry out daily functional activities which they would encounter every day.

Marsh and Redmayne investigated the correlation between self-concept and the field testing of physical fitness. The Physical Ability scale was correlated with all the specific components of physical self-concept. Self-concepts of endurance ($r = .64$), strength ($r = .59$), and flexibility ($r = .49$) gained the strongest positive correlation with the endurance from fitness test. Other items did not show satisfactory correlation. One of the reasons for that, as commented by the authors, could be due to the small and unrepresentative sample of young adolescent girls. Girls at that age are experiencing many physiological and psychological changes and these might have caused discrepancies. Another reason would be the limited choice of fitness tests (Marsh & Redmayne, 1994). In brief, these two self-reported scales failed to establish satisfactory correlations and validity between the questionnaire and the physical fitness assessed from any form of fitness test.

Table 2

Concurrent validity of each factor in older and younger adult groups of the Perceived Physical Fitness Scale (Abadie, 1988).

	Older (> 50 yr., M=62.3) (n=67)		Younger (< 50 yr., M=27.8) (n=28)	
Factors	Criterion	r	Criterion	r
Physical condition	Symptoms-limited maximal exercise stress test	.43**	VO2max estimation from Astrand-Rhyming test	.61**
Muscular strength	Hand-grip dynamometer	.07*	Hoeger's bench press test	.47**
Muscular flexibility	Modified sit-and-reach	.12*	Performance on modified sit-and-reach	.53**
Body composition	Skinfolds (Jackson & Pollock, 1978)	-.67**	Skinfolds thickness (Jackson & Pollock, 1978)	-.68**

Note. $^{*}p > .05$. $^{**}p < .01$.

Merits and Limitations of Self-Reported Questionnaires

Like laboratory and field testing, self-reported questionnaires bear merits and limitations at the same time. Feasible self-administered questionnaires could be found in previous studies. A self-administered questionnaire for health risk appraisal was tested on its feasibility and acceptability (Stuck et al., 2002). The participants who completed the questionnaire were heterogeneous older people from three

European countries. The results showed that the instrument was highly accepted and was able to identify modifiable risk factors for functional status decline. Another self-rated health inventory set questions the ADLs of people aged 65 or above (Gama et al., 2000). Questions are as basic as "use of stairs" and "dressing", and answers are to be chosen from a five-point scale. The association between the self-rated health and ADLs was high. The two highest activities were use of stairs (odds ratio = 4.28, 95% CI = 2.82 - 6.52) and ambulation (odds ratio = 3.67, 95% CI = 2.39 - 5.64).

However, not everything can be assessed with a self-reported questionnaire. For instance, a previous study showed that the self-reported body weight of older women should not be relied on if individual accuracy is required in research and clinical practice (Lawlor et al., 2002). If there is any missing value in the questionnaire, it could be improved by changing the questionnaire layout or by having the interview better administrated (Bjorner & Kristensen, 1999).

A self-reported questionnaire is convenient in the way that no monitor is required (i.e., self-administered). It can be done at any time, in the most convenient time, for the person, which means the labor cost is minimized. This is especially true for the always busy Hong Kong people who are usually from lower socioeconomic status (SES) groups. However, from similar past experience of self-reported questionnaires in other fields, it is believed that a self-reported questionnaire to predict physical fitness is feasible.

Even an imperfect implementation of self-assessment could generate benefits, such as drawing awareness of the concerned issues and suggesting any strategy for improvement (Teo & Dale, 1997). A self-reported or self-rated questionnaire could help the organizers work out where the problem lies among the participants (Gama et al., 2000).

Setting of a Questionnaire

It has been suggested that when measuring the intensity of physical activity, "moderately accurate predictions of fitness can be made from age, body mass, skinfold thicknesses, and global assessments of habitual physical activity without engaging participants in an exercise test" (Shephard, 2003, p.200). If this is viable, it may be possible to create an inventory by including these items to estimate an individual's physical fitness.

As the format of a questionnaire could affect a person's intention to complete it, it is important to have the format carefully considered. The format includes the length of survey, order of questions, layout (Jackson & Furnham, 2000), the choice of a rating scale (Siegle, n.d.; Trochim, 2005), and wording (Heberlein & Baumgartner, 1978). Response rates could also be affected by editorial format such as the cover design, an interesting title, and the size and style of font, which should also be considered (Sletto, 1940). Whether the number of rating is to be a single-point or an odd-point would also have effect on the response (Kalton et al., 1980).

Pre-testing is urged before the final questionnaire is sent out (Sletto, 1940). According to Sletto, the purpose of pre-testing a questionnaire is to obtain tentative answers to many problems during the developing stage of a questionnaire, for instance, knowing how the attractiveness could be enhanced. Through trial-and-error, the final version can increase the response rate and moderate possible criticism and condemnation.

Summary

Physical fitness consists of five components. Because of its impact on health, everyone should be aware of his/her level of physical fitness. Existing methods to assess an individual's fitness include laboratory and field tests. The former usually delivers a more accurate result and is often referred to as the gold standard, but it relies on the availability of sophisticated equipment, considerable time, high level of skills, and manpower. The latter one is a more simple form of assessment. It takes less time, requires less skill, is more easily manageable, and is more convenient. Compared with a laboratory test, a field test is more widely used, especially when managing a large group of people. Nevertheless, both laboratory test and field tests require considerable time and trained personnel, and are not suitable for mass testing that examines thousand of participants within a relatively short period of time. It is therefore valuable to develop a health-related played fitness assessment using a self-reported questionnaire.

Based on personal experiences, developing a questionnaire for assessing the five-component physical fitness might also be practicable. Such a questionnaire would be a very useful tool in conducting epidemiological studies. Based on this rationale, this study would attempt to develop a self-reported questionnaire for assessing health-related physical fitness.

CHAPTER III

METHODOLOGY

The objective of this study was to develop a self-reported questionnaire that would allow people to evaluate their physical fitness. In order to do so, question items were constructed in the first stage according to the five components of health-related physical fitness. The content validity was examined by five experts in the field of physical fitness, followed by a pilot test for evaluating its readability, comprehension, and administration of the questionnaire. Subsequently, criterion-related validity was examined by recruiting a sample of Chinese adults who completed both the questionnaire twice and a series of fitness tests in a laboratory. Test-retest reliability and internal consistency of the questionnaire were examined.

Development of Questionnaire

Item Construction

The items were constructed by the investigator in Chinese (Appendix A) through observations and investigations on functional fitness activities in daily life routines, such as walking upstairs in the subway (cardiorespiratory capacity) or lifting the goods available in a supermarket (muscular strength or endurance). An attempt was made to write as many question items as possible for each health-related fitness component. The SPFQ was originally written in Chinese (Appendix B). An English version was prepared for reference and is shown in Appendix C. Each question described an experience in daily living that involves functional abilities that are related to fitness. Respondents were asked to rate their perceived physical proficiency in performing these functional activities described in the questionnaire items on a five-point Likert scale. They were expected to rate from 1 to 5 points. Reversed scoring was randomly applied on some items to prevent the respondents from cheating. For question items of cardiorespiratory fitness, functional abilities that related to cardiorespiratory fitness such as walking, jogging, or running were asked. For items related to body composition, physical movements that could be affected by physique were asked. For items of muscular strength, muscular tasks such as lifting goods and furniture were asked. For items of muscular endurance, repeated muscular activities such as carrying objects of different weights and sizes continuously were asked. For items related to flexibility, physical movements or activities that involve joint mobility (limbs and back) were asked. As a result, assembling items describing cardiorespiratory capacity (17 items), muscular strength (18 items), muscular endurance (16 items), flexibility (17 items), and body composition (16 items), a total of 84 items, were constructed in the initial stage.

Content Validity

All question items constructed in this initial stage were reviewed by five experts in physical fitness for examining content validity. Experts were requested to review the relevancy and appropriateness of each question items one by one, and to determine if the question items were capable of evaluating the physical fitness of respondents. For each question item being reviewed, three options were provided: (a) Yes; (b) Yes, but needs modification; and (c) No. Only items receiving 80% or above of agreement from the five experts were retained (i.e., at least four experts out of five chose options A or B). Items were revised accordingly for items that required modification as suggested by reviewers. As a result, 77 out of the original 84 items were retained in the SPFQ.

Pilot Test

A small convenience sample of 10 adults from a university campus was involved in the pilot test. They were requested to review the readability, comprehensibility, and administration of the questionnaire. Further revisions were made according to their suggestions.

Validity and Reliability

Participants. A sample of 169 Chinese participants (83 women, 86 men, aged 24-59 yrs) was recruited through local newspaper advertisements, the university's homepage, and alumni newsletters. To evaluate the validity of the SPFQ in Chinese adults, participants needed to engage in laboratory fitness tests which involved measurements of HR, percent body fat, muscular functions, and flexibility. The purpose was to ensure normal physiological responses. Only participants who met the following inclusion criteria were included:

(1) Chinese adults (male and female, aged 20-60 yrs);

(2) no known diseases as screened by a Health History Questionnaire (Appendix D);

(3) not receiving any medication that may affect their physical function, such as drugs for hypertension or high cholesterol, regardless of whether or not they had the disease at the time of this study; and

(4) screened by the PAR-Q (Appendix E).

All eligible participants were informed through telephone calls prior to the day of testing to maintain their normal diet habits and not to drink any caffeinated drinks (e.g., coffee and tea) before the laboratory fitness test. They were also asked to wear proper exercise clothing and sports shoes for the fitness test. Procedures in this study were also given to them: to take part in a series of physical fitness test, as well as to complete the SPFQ twice on separate days.

Procedure. Upon arrival at the laboratory, participants were asked to complete the informed consent form (Appendix F), to self-report their health history, and to return the SPFQ (Trial 1) and the completed PARQ that they were deemed able to participate in physical activity safely. The participants were then briefed about the procedures of the fitness tests. Before taking part in the fitness tests, resting measurements such as heart rate (HR) and blood pressure (both with the Panasonic Blood Pressure Meter, EW3100), weight in kilograms (on the electronic Tanita Body Fat Analyzer, TBF-401, Tokyo, Japan) and height in centimeters were recorded as baseline. Participants were required to take off their shoes and socks and to wear light exercise clothing during these measurements. Skinfold measurement was also taken with a calibrated Harpenden caliper (England) before starting any exercise test.

Each participant was instructed to perform 5 to 10 minutes of warm-up stretching before the exercise tests of isometric and handgrip tests, sit-and-reach tests, and curl-up test. The VO_2max on the treadmill and the cool-down were the last two stages. The participant was released only when he/she reported no physical discomfort after the exercises. He/she was given the SPFQ (Trial 2) before leaving, and was asked to return by fax the completed questionnaire after about one week's time.

All fitness tests were carried out by trained and experienced fitness testers at the laboratories in the Department of Sports Science and Physical Education, the Chinese University of Hong Kong.

Criterion Measures

Cardiorespiratory fitness. Estimated VO_2max measured from a multi-staged Bruce treadmill test was determined as a criterion of cardiorespiratory fitness (Bruce et al., 1973). According to the Bruce protocol, five stages were adopted which started with a 2-minute warm-up and ended with a 5-minute recovery on the treadmill (SCIFIT AC7000, Tulsa, USA). In the warm-up stage, speed was 1.7 mph and slope was 0%. For the first stage, speed was 1.7 mph and the gradient was 10%. Then both speed and elevation were increased every three minutes for the subsequent stages. The test was terminated when the exercise heart rate reached 85% of the participants' maximal heart rate [(220 – age) x 85%], or when he/she felt discomfort. When the test finished, all participants were required to walk on the treadmill slowly for another five minutes for recovery.

The VO_2 of the participant was measured and recorded with a COSMED portable metabolic measuring system ($K4b^2$system, Rome, Italy). The system was calibrated before the exercise test on the day of testing with reference to the manual to ensure that the oxygen and carbon dioxide were measured accurately. HR was measured by a POLAR heart rate monitor (Tempo, Kempele, Finland). Throughout the test, participants were required to wear a facemask connected to a turbine and a sampling line, and to wear the $K4b^2$ system on their body. VO_2max was predicted with a double-point equation (Heyward, 2002, p.67):

$$VO_{2max} = VO_{2\text{-}85} + \left(\frac{VO_{2\text{-}85} - VO_{2\text{-}60}}{HR_{85} - HR_{60}} \right) \times (HR_{max} - HR_{85})$$

where

$VO_{2\text{-}85}$ and HR_{85} = VO_2 and HR measured at 85% of maximum HR, respectively;

$VO_{2\text{-}60}$ and HR_{60} = VO_2 and HR measured at 60% of maximum HR, respectively;

and

HR_{max} = 220 – age.

The estimated VO_2max was then correlated to the score of cardiorespiratory component measured from the SPFQ.

Body composition. For the component of body composition, skinfold measurement was obtained as the criterion. Definitions of skinfold sites and measurement techniques were adopted from the ACSM's Guidelines (2006). Six (women) or seven (men) skinfold sites were taken. These sites included triceps, biceps, chest (men only), abdomen, suprailiac, thigh, and calf. Each site was measured two or three times by the same fitness tester and the final reading was the average of the two nearest readings. All measurements were done on the right side of the participant. Standard description of skinfold sites and procedures are described in the ACSM's Guidelines (2006). Skinfold measurement of these sites was summed to correlate with the score of body composition component measured from the SPFQ.

Muscular strength. The criterion of muscular strength was determined from isometric strength tests using a handgrip (Takei Scientific Instruments Co. Ltd, T.K.K. 5101 Grip D, Tokyo, Japan) and isometric dynamometers (Takei Scientific Instruments Co. Ltd, T.K.K. 5102 Back D, Tokyo, Japan). The testing procedures and guidelines were explained in detail elsewhere (Baumgartner & Jackson, 1999, pp. 202-211). The fitness tester demonstrated the correct way to perform the tests. Each participant was asked to perform with 50% of his/her physical effort as familiarization and preparation procedures. The participant was then given three trials at maximum effort for the handgrip test and two trials for other lifting tests. The sum of averaged handgrip strength of each hand and the averaged strength of each lift (arm, shoulder, back, and leg) were obtained from the strength tests to correlate with the score of the muscular strength component measured from the SPFQ.

Muscular endurance. The criterion of muscular endurance was determined by a curl-up test. The participants were asked to lie on a mat in the laboratory, bend their knees, and repeatedly lift their upper body using the abdominal strength until the middle of their palms could touch the highest point of the bent knees. Arms should also be straight. The participants were to do as many repetitions as possible at the cadence of 50 rpm. Procedures for the measurement are published elsewhere (ACSM, 2006). Before the test, the participants had to perform one or two curl-ups according to the fitness tester's instruction to familiarize themselves and to prepare for the exercise. The curling up and lying down of the upper body were counted as one curl-up. Only one trial was administered. If the participant could not demonstrate a correct curl-up, he/she was asked to stop and was regarded as making 0 curl-ups in this test; otherwise, the test was terminated once the participant could not follow the cadence or the posture became incorrect. The number of times of curl-up was then correlated to the score of muscular endurance component measured from the SPFQ.

Flexibility. The criterion of flexibility was determined by the CUHK sit-and-reach test on a bench (Hui & Yuen, 2000). Measurement was made on one leg three times, then the other leg another three times. Participants were required to take off their shoes for this test. Procedures were described in the study of Hui and Yuen. The final score of flexibility was the sum of the best measurement obtained from each leg in centimeters.

Statistical Analysis

Descriptive statistics

Means and standard deviations were calculated for all anthropometric characteristics and physical fitness measurements.

Reliability

The test-retest reliability was analyzed by the Pearson Correlation between scores from Trial 1 and Trial 2 for women, men, and all participants. Each of the 77 items from Trial 1 was correlated to the corresponding item from Trial 2.

Completed Questionnaires were used for examining the reliability of internal consistency. Cronbach alpha reliability (α) among all items within each fitness component was computed once for each component in each of the two trials. Alpha reliability for women, men, and all participants was also computed.

Criterion-related validity.

The criterion-related validity of the SPFQ was determined by computing the Pearson correlations between the scores obtained from the fitness tests and the corresponding component scores from the SPFQ, adjusted for women and men for both trials of the SPFQ. The correlation coefficient of each question item within the same component to the corresponding criterion was examined one by one. Items with a lower correlation to the criterion tests were eliminated. The criterion-related validity of the remaining items was examined again. All statistical analyses were performed with the SPSS software, Version 12.0.

CHAPTER IV

RESULTS

Descriptive Statistics

Descriptive statistics of participants and the criterion fitness test results are presented in Table 3. Since six women and 14 men failed to complete the second questionnaire (Trial 2), the sample size for the test-retest reliability (n = 149) was smaller than other analyses (n = 169). As can be seen from the descriptive statistics, the physical fitness profiles of the sample were considered average to good as compared to the norms reported for Chinese of similar age in Hong Kong (許世全, 2002).

Descriptive statistics of participants and their average scores in each component in the SPFQ of women and men are reported in Tables 4 and 5, respectively. The mean scores from Trials 1 and 2 are given in each table. From these two tables, it was noted that men generally rated higher than women in all fitness components. The rating of men's muscular strength and women's flexibility were higher than other physical fitness components in their respective gender group.

Table 3

Descriptive statistics of the participants and laboratory criterion fitness tests.

	Women		Men	
Variable	N	Mean ± SD	n	Mean ± SD
Age (years)	83	46.66 ± 7.08	86	47.16 ± 6.60
VO_{2max} (ml/kg/min)	83	31.87 ± 4.78	85	36.01 ± 7.32
BMI (kg/m^2)	83	22.83 ± 2.93	86	24.23 ± 3.12
WHR	83	0.77 ± 0.05	86	0.88 ± 0.05
Body Fat (%)	83	29.13 ± 5.58	86	22.36 ± 5.13
Sum of skinfold (mm)	83	124.41 ± 30.25	85	113.89 ± 35.56
Isometric strength (kg)	81	145.90 ± 31.54	83	232.35 ± 54.63
Curl-up test (rep)	83	6.658 ± 9.39	85	15.29 ± 11.11
Sit-and-reach test (cm)	82	104.70 ± 24.36	86	89.90 ± 24.79

Note. WHR = waist-to-hip ratio, measured in cm; Body Fat in percentage of body weight measured from Bioelectrical Impedance Analysis (BIA); Sit-and-reach test = sum of left and right legs.

Table 4

Scores of SPFQ (women).

Variable	n	Trial 1 (Mean ± SD)	n	Trial 2 (Mean ± SD)
CR	78	3.11 ± 0.65	77	3.10 ± 0.74
BC	78	3.66 ± 0.64	77	3.64 ± 0.68
MS	79	3.44 ± 0.76	74	3.50 ± 0.79
ME	80	3.35 ± 0.79	76	3.37 ± 0.87
FLEX	78	3.80 ± 0.59	76	3.79 ± 0.66

Note. CR = scores from the cardio-respiratory endurance component; BC = scores from the body composition component; MS = scores from the muscular strength component; ME = scores from the muscular endurance component; FLEX = scores from the flexibility component.

Table 5

Scores of SPFQ (men).

Variable	n	Trial 1 (Mean ± SD)	n	Trial 2 (Mean ± SD)
CR	84	3.47 ± 0.59	72	3.42 ± 0.65
BC	85	3.85 ± 0.58	71	3.89 ± 0.54
MS	84	4.06 ± 0.59	72	4.05 ± 0.64
ME	83	4.05 ± 0.66	72	4.05 ± 0.71
FLEX	81	3.94 ± 0.59	67	3.97 ± 0.66

Note. CR = scores from the cardio-respiratory endurance component; BC = scores from the body composition component; MS = scores from the muscular strength component; ME = scores from the muscular endurance component; FLEX = scores from the flexibility component.

Criterion-related Validity

The criterion-related validity of women and men were analyzed for Trial 1 and Trial 2 separately. As can be seen from Table 6, for each of the fitness components of women, there was very little difference in the criterion-related validity between Trial 1 and Trial 2. Similar results were found for men. The criterion-related validities for CR in both Trial 1 and Trial 2 were close to zero.

The ME component had low correlation with the criterion test ($r = .07 - .16$), while FELX ($r = .23 - .29$) and MS components ($r = .20 - .28$) were slightly higher. The BC component was the only one that yielded a moderate correlation with the criterion tests in men [$r = (-.44) - (-.42)$] and women [$r = (-.50) - (-.53)$].

Because the result of the examination of criterion-related validity of the SPFQ scores was not satisfactory, an attempt was made to remove items that had low correlations with the criterion to improve the validity. It was hoped that a better validity could be achieved through the revised questionnaire.

Table 6

Criterion-related validity of SPFQ in Trial 1 and Trial 2.

		Women	Men
CR	Trial 1	.05	- .08
	Trial 2	.10	- .02
BC	Trial 1	- .50**	- .44**
	Trial 2	- .53**	- .42**
MS	Trial 1	.27*	.20
	Trial 2	.21	.20
ME	Trial 1	.12	.16
	Trial 2	.07	.15
FLEX	Trial 1	.24*	.29*
	Trial 2	.23*	.24*

*$p < .05$. ** $p < .01$

Note. CR = scores from the cardio-respiratory endurance component; BC = scores from the body composition component; MS = scores from the muscular strength component; ME = scores from the muscular endurance component; FLEX = scores from the flexibility component.

Reliability

Test-retest Reliability

Results from the Pearson correlation between scores from Trial 1 and Trial 2 for women, men, and all participants are shown in Table 7. Each of the 77 items from Trial 1 was correlated to the corresponding item from Trial 2. In Trial 2, 6 women and 14 men did not complete the questionnaire and therefore only responses from the remaining 149 participants (77 women, 72 men) were examined for the test-retest reliability analysis. Very high test-retest reliabilities for women and men for all components ($r = .84 - .92$) were obtained. Similar results were derived when data from women and men was combined.

Table 7

Test-retest reliabilities (Pearson r) for women, men, and all participants.

	Women	Men	All
CR	.91**	.87**	.90**
BC	.92**	.84**	.89**
MS	.92**	.88**	.92**
ME	.92**	.87**	.92**
FLEX	.85**	.89**	.87**

** $p < .01$.

Note. CR = scores from the cardio-respiratory endurance component; BC = scores from the body composition component; MS = scores from the muscular strength component; ME = scores from the muscular endurance component; FLEX = scores from the flexibility component.

Internal Consistency

Internal consistency scores were calculated separately for women and men and are presented in Table 8. Missing values in the questionnaires were excluded from the analysis.

The Cronbach's Alpha values for women and men revealed that the SPFQ scores are highly reliable in terms of internal consistency (all $\alpha > .85$). As shown in Table 8, the MS component score yielded the highest Alpha values ($\alpha = .90 – .94$), and BC yielded the lowest ($\alpha = .85 – .89$) in women, men, and all; even so, the overall internal consistency was still very high.

Table 8

Internal consistency for women, men, and all participants.

	Women		Men		All	
	Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2
CR	.88	.93	.87	.90	.89	.92
BC	.88	.89	.85	.86	.87	.88
MS	.92	.94	.90	.92	.92	.94
ME	.90	.93	.88	.91	.91	.94
FLEX	.87	.92	.87	.90	.87	.91

Note. CR = scores from the cardio-respiratory endurance component; BC = scores from the body composition component; MS = scores from the muscular strength component; ME = scores from the muscular endurance component; FLEX = scores from the flexibility component.

Questionnaire Revision

The criterion-related validity of each item was generated with SPSS. Items with high correlations with the corresponding criterion in Trial 1 were selected. Since the criterion-related validity of all CR items approached zero for women and men, the CR component was not considered in this questionnaire revision. For each of the other four fitness components, two combinations were created for women (Table 9) and men (Table 10). Two versions of item reduction strategies were adopted as follows:

Revision I. The selection criteria were randomly independent for each component. The number of selected items of each component differed from each other, and differed between women and men. A summary of the selection criteria is presented in Tables 9 and 10 for women and men, respectively. As a result, for women, the number of items retained for each component were: 6 out of 16 items from BC, 8 out of 17 items from MS, 4 out of 12 items from ME, and 5 out of 17 items from FLEX; and for men, 5 out of 16 items from BC, 4 out of 17 items from MS, 3 out of 12 items from ME, and 5 out of 17 items from FLEX. With this strategy, every component of women and men had significant improvement in the criterion-related validity, the result of which is summarized in Table 11.

Table 9

Selected items from Trial 1 (women).

	Revision I		Revision II	
	Selection criteria	Items retained	Selection criteria	Items retained
BC	$r \geq .30$	4, 45, 49, 50, 55, 66	$r \geq .40$	4, 45, 50, 55
MS	$r \geq .20$	3, 28, 32, 34, 36, 43, 63, 76	$r \geq .22$	3, 28, 34, 36, 76
ME	$r \geq .13$	16, 53, 59, 64	$r \geq .15$	17, 60
FLEX	$r \geq .24$	6, 31, 48, 73, 75	$r \geq .25$	6, 31, 48, 75

Note. BC = scores from the body composition component; MS = scores from the muscular strength component; ME = scores from the muscular endurance component; FLEX = scores from the flexibility component.

Table 10

Selected items from Trial 1 (men).

	Revision I		Revision II	
	Selection criteria	Items retained	Selection criteria	Items retained
BC	$r \geq .30$	4, 7, 50, 55, 66	$r \geq .40$	4, 50, 55, 66
MS	$r \geq .20$	1, 3, 8, 34	$r \geq .25$	3, 8
ME	$r \geq .15$	46, 64, 68	$r \geq .25$	64
FLEX	$r \geq .25$	18, 35, 47, 71, 77	$r \geq .36$	18, 71

Note. BC = scores from the body composition component; MS = scores from the muscular strength component; ME = scores from the muscular endurance component; FLEX = scores from the flexibility component.

Revision II. The purpose of having Revision II was to achieve an even higher criterion-related validity with a more stringent selection criterion; thus, fewer items for every component. The selection criteria were again random and independent for each component. Consequently, for women, there were 4 out of 16 items from BC, 5 out of 17 items from MS, 2 out of 12 items from ME, and 4 out of 17 items from FLEX; and for men, there were 4 out of 16 components from BC, 2 out of 17 items from MS, 1 out of 12 items from ME, and 2 out of 17 items from FLEX. It was found that in applying a more stringent selection criterion, the criterion-related validity of these four components increased in women and men, except for the men's MS component (Table 11).

Table 11 compares the criterion-related validity between the original SPFQ and the two revisions of selected items. In Revision II, the validity further increased for all components, except men's MS. All validity coefficients were significant (p <.05), except women's ME. All criterion-related validity coefficients in both revisions were r > .20. However, although fewer items were included in Revision II, the validities were higher than the original and Revision I. The validity coefficients generated from Revisions 1 and 2 were very similar.

Table 11

Comparison of criterion-related validity of the SPFQ in original and revised questionnaire (Revisions 1 and 2).

	Women			Men		
	Original	Revision I	Revision II	Original	Revision I	Revision II
BC	- .50**	-.57**	-.60**	-.44**	-.57**	-.57**
MS	.27*	.30**	.30**	.20	.32**	.31**
ME	.12	.20	.20	.16	.21	.25*
FLEX	.24*	.38**	.39**	.29*	.50**	.57**

* p < .05; ** p < .01

Note. BC = scores from the body composition component; MS = scores from the muscular strength component; ME = scores from the muscular endurance component; FLEX = scores from the flexibility component.

The two revised groups of items later had their criterion-related validity reexamined in Trial 2. Items selected from Trial 1 were then turned into a new questionnaire to test for the criterion-related validity in Trial 2 and to test whether similar results could be generated in Trial 2. Results of women and men are presented in Tables 12 and 13, respectively.

For women, except MS and ME components, the other two components (BC and FLEX) had similar criterion-related validity (Table 12). When applied to the selected question items from Trial 1 to Trial 2, the correlation to the criteria of MS and ME components (in both revisions) were lower than that in the original SPFQ; only that of the BC component were the highest in Trial 2. The validity coefficients of FLEX component in Trial 2 were higher than the original but lower than Trial 1 (Table 12).

For men, only the correlation coefficient of ME Revision I in Trial 1 was lower than that in the original SPFQ (Table 13). Despite the slight drops of all components in Trial 2, other validity coefficients showed that the criterion-related validities of other components in Trial 2 were rather similar to that in Trial 1.

When the revised items (both Revisions 1 and 2) derived from Trial 1 were tested in Trial 2 for validities, the above observation was not true in women's MS and ME components (Table 12) and men's ME Revision I components (Table 13). Only BC and FLEX components had improvements in the validity in both revisions and both trials.

In conclusion, generally speaking, the criterion-related validity of Revision II was higher than the original SPFQ and Revision I, with the exception of MS and ME components in women and ME in men. When considering a more practical SPFQ in terms of higher criterion-related validities and fewer items included, the Revision II of SPFQ is hence recommended.

Table 12

Comparison of criterion-related validity of Trials 1 and 2 of SPFQ in original and two revisions of revised questionnaire (women).

	Original	Trial 1		Trial 2	
		Revision I	Revision II	Revision I	Revision II
BC	- .50**	-.57**	-.60**	-.60**	-.66**
MS	.27*	.30**	.30**	.20	.23*
ME	.12	.20	.20	.11	.07
FLEX	.24*	.38**	.39**	.25*	.27*

* $p < .05$; ** $p < .01$

Note. BC = scores from the body composition component; MS = scores from the muscular strength component; ME = scores from the muscular endurance component; FLEX = scores from the flexibility component.

Table 13

Comparison of criterion-related validity of Trials 1 and 2 of SPFQ in original and two revisions of revised questionnaire (men).

	Original	Trial 1		Trial 2	
		Revision I	Revision II	Revision I	Revision II
BC	-.44**	-.57**	-.57**	-.53**	-.53**
MS	.20	.32**	.31**	.29*	.26*
ME	.16	.21	.25*	.12	.20
FLEX	.29*	.50**	.57**	.43**	.48**

* $p < .05$; ** $p < .01$

Note. BC = scores from the body composition component; MS = scores from the muscular strength component; ME = scores from the muscular endurance component; FLEX = scores from the flexibility component.

CHAPTER V

DISCUSSION

The SPFQ was developed to evaluate health-related physical fitness level among Hong Kong Chinese adults in the form of self-administration. The question items were designed to evaluate daily functional activities. Results showed a very high test-retest reliability ($r = .84 - .92$) and internal consistency ($\alpha > .85$) for all of the five components. However, the criterion-related validity of all components was low, except for the BC component which yielded a moderate validity. When items with low correlation to the respective criterion were removed, criterion-related validity increased substantially except for the CR question items. When revising the questionnaire, items of the CR component were not included because the correlations of all CR items with its criterion approached zero. The revised questionnaire, therefore, consisted of items of the BC, MS, ME, and FLEX components only. The total number of items included in the final version of SPFQ between women (15 items) and men (9 items) was different. The revised SPFQ for women and men are presented in Appendixes G and H, respectively.

Reliability

Test-retest Reliability

Although limited studies investigated the reliability of a self-reported physical fitness questionnaire, a number of studies were found investigating the reliability of a self-administered physical activity questionnaire. A summary of these studies are shown in Table 14.

Table 14

Comparison of test-retest reliability of self-report physical fitness/ physical activity questionnaires.

Studies	Variable	Reliability (r)
SPFQ	Fitness:	
(Present study)	Cardiorespiratory	.90**
	Body composition	.89**
	Muscular strength	.92**
	Muscular endurance	.92**
	Flexibility	.87**
Perceived Physical Fitness	Fitness:	.92
Scale	Physical condition	--
(Abadie, 1988)	Muscular flexibility	--
	Muscular condition	--
	Body composition	--
Baecke Questionnaire of	Work index	.88
Habitual Physical Activity	Sport index	.81
(Baecke et al., 1982)	Leisure index	.74
GLTEQ	Leisure-time exercise:	
(Godin & Shephard, 1985)	Light	.48*
	Moderate	.46*
	Strenuous	.94*
	Total	.74*
	Sweat	.80*

Table 14 (Continued)

Studies	Variable	Reliability (r)
PSDQ	Nine physical self-concepts	.83 (3-mo);
(Marsh et al., 1996)		.69 (14-mo)
TOQ	Occupational activities:	
(Ainsworth et al., 1961)	Activity units	.83*
	H/ wk	.63
	Average METs	.37

* $p < .05$. ** $p < .01$

Note. SPFQ = Self-reported Physical Fitness Questionnaire; GLTEQ = Godin Leisure-Time Exercise Questionnaire; PSDQ = Physical Self-Description Questionnaire; TOQ = Tecumseh Occupational Physical Activity Questionnaire.

In Abadie's study of Perceived Physical Fitness Scale, a high test-retest reliability coefficient of all items was reported ($r = .92$); the reliability coefficients of other independent components, however, were not reported. In Baecke Questionnaire of Habitual Physical Activity, test-retest reliability coefficients in three different occasions (work activity, during sport, and non-sports leisure activity) were moderately high ($r \geq .74$) (Baecke et al., 1982). The four-item Godin Leisure-Time Exercise Questionnaire (GLTEQ), which is a self-explanatory questionnaire measuring usual leisure-time exercise behavior, reported fair ($r = .48$, light activities) to high ($r = .94$, strenuous activities) test-retest reliability (Godin & Shephard, 1985). The one-month test-retest reliability coefficients for the self-explanatory Tecumseh Occupational Questionnaire (TOQ) were fair to moderately high for three different units of physical activity measurements (Ainsworth et al., 1993).

Marsh developed the Physical Self-Description Questionnaire (PSDQ) to assess physical self-concept (Marsh et al., 1994). Similar to the SPFQ in the present study, it contains 70 items and 11 scales for measuring self-concept, in which only four scales are related to health, however. The PSDQ uses a six-point true/false response scale while the SPFQ uses a five-point Likert Scale. The test-retest reliability coefficients of .83 (3-month period) and .69 (14-month period), and an internal consistency of .92 was reported (Marsh, 1996). However the PSDQ includes scales in measures not related to health-related fitness components, such as coordination and sports competence.

In the present study, the test-retest reliability coefficients of SPFQ were very high for both women and men. A similar result was obtained when the data of women and men were combined for analysis. The test-retest reliability coefficients were higher than the Baecke Questionnaire, the GLTEQ, and the TOQ. However, the latter three questionnaires measure physical behavior while SPFQ measures physiological aspect of functional fitness. Although the PSDQ was designed to assess physical self-concept, which has a similar direction of the SPFQ, the test-retest reliability was not as high as the SPFQ. The test-retest reliability of the SPFQ was comparable to the study of Abadie's questionnaire. However, the reliability of each component in Abadie's study was not reported. It is also important to note that Abadie's Perceived Physical Fitness Scale measures the perception of physical fitness instead of actual functional ability in performing daily activities as in SPFQ.

Internal Consistency

Limited studies reported the internal consistency of self-report questionnaires. Of the above five questionnaire studies, only the PSDQ had demonstrated internal consistencies. A comparison between the internal consistency of PSDQ and SPFQ is displayed in Table 15. The PSDQ was developed by Marsh in 1994, but the internal consistency was not examined until a recent study of a 47-item short version of

PSDQ (Peart et al., 2007). The alpha coefficients among those four health-related scales of physical self-concept in that short version of PSDQ ($\alpha = .88 - .92$) were very similar to the alpha coefficients among the five fitness components in the SPFQ of the present study ($\alpha = .87 - .92$).

Table 15

Comparison of internal consistency in SPFQ and PSDQ.

Studies	Variable	α
SPFQ	Fitness:	
(Present study)	Cardiorespiratory	.89
	Body composition	.87
	Muscular strength	.92
	Muscular endurance	.91
	Flexibility	.87
PSDQ	Physical self-concept:	
(Peart et al., 2007)	Body Fat	.90
	Endurance/ fitness	.89
	Flexibility	.88
	Strength	.92

Note. SPFQ = Self-reported Physical Fitness Questionnaire; PSDQ = Physical Self-Description Questionnaire.

Compared with other studies, the SPFQ of the present study is reliable and comparable to other established questionnaires in terms of its test-retest reliability and internal consistency.

Validity

For the SPFQ, the criterion-related validity coefficients before a questionnaire revision of all components were above .10, except for the component of CR. The validity coefficients from the CR component were very close to zero for both women and men, which implied that the CR question items had almost no correlation with the measured VO_2max.

To improve the correlation of the components with the corresponding criterion tests, the SPFQ had been revised. Some items were removed according to their low correlation with the corresponding criterion. Nevertheless, the validity of CR did not improve much; it was still poorly correlated with the criterion of VO_2max even when some items were removed.

The failure of the CR items to measure cardiovascular fitness may be due to the way that the questions were asked and interpreted. There were four major types of CR questions in SPFQ: (1) relating to walking, jogging, and running different distances in different settings, such as running at a medium speed for 20 minutes, jogging for 20 minutes, and running 400 meters at a medium speed; (2) physical reaction after kneeling or squatting on the floor, such as feeling dizzy, breathing heavily, and heavy heartbeats; (3) walking or jogging pace compared with others at a similar age; and (4) their physical reaction of walking or running in daily life experiences. Distances, number of steps, time, and physical tasks being described in the question items were specified. However, the participants may still have difficulties to perceive or to realize these descriptions such as how long a 50-meter distance or a standard athletic field is, or how many steps of staircases they could have handled in the real life situation. On the other hand, because the participants possess different lifestyles and life experiences, some of them may not have substantial experiences in traveling on the bus or the train. Also, some of them may never be aware of their own and others' walking or jogging pace or habits. It would therefore be difficult for them to answer questions which made comparisons with others.

When the participants encountered such question items as above, different daily lifestyles and experiences and lack of awareness and experiences may be the reasons that led to the unsatisfactory validity of CR component.

Another possible reason for the low validity of the CR component may due to the inaccuracy of the estimated VO_2max from the laboratory testing. In the present study, the VO_2max was estimated from a submaximal test on the treadmill. Availability of measured VO_2max from maximal test, which is accepted as the criterion measure, may provide a better result for validation. However, because of the high costs associated with equipment, space, and manpower required in performing these tests, it is not a desirable method in the present research setting (ACSM, 2000, p. 68-70).

Even though the validity coefficients of all other components improved after deleting some items, the validity coefficients were considered low to moderate (women: $r = .20\text{-}(-.60)$; men: $r = .25\text{-}(-.57)$). The criterion-related validity of the SPFQ was shown to be fair to moderate in other four fitness components for both women and men. The coefficients were very similar between genders, except for FLEX. Among the four components, the SPFQ yielded better validity in BC and FLEX than in the other two components. This result was also similar both in women and men.

The reasons that BC scores showed higher correlations with objective measures than other scores may due to the content of the questions. Unlike questions asked for other components, questions produced for BC were very simple and were of basic bodily movements or life experiences that everyone would come across. For example, if a person had high body fat content, when standing with legs together, it would thus be likely that the inner portions of his/her feet cannot touch each other (SPFQ No. 19). It was probably because whether a person is fat or not can be told by his/her physical appearance, or can even be sensed by the person. Therefore, the high BC scores may not imply that people were more aware of this aspect of their fitness, but may imply that people may approximately know whether they are fit or fat through some of their daily activities and life experiences.

Although some studies seemed to present higher validity coefficients than SPFQ, none of the analysis was adjusted by gender. As can be seen from Table 16, different studies estimated validity in different ways. The criterion-related validity coefficients obtained for the four fitness components from the revised SPFQ were $r = .20 - .60$ for women and $r = .25 - .57$ for men. In Abadie's study, validity coefficients were reported in two age groups, namely, older ($\geq$ 50 years) and younger (21 - 50 years) age groups (Abadie, 1988). In the study of GLTEQ, validity coefficients of physical activity were reported according to the types of criterion (VO_2max and percent body fat (BF)) (Godin & Shephard, 1985). In the study of TOQ, the relationship between activity units and activity questionnaires and various physiological measures (age and gender-adjusted) was reported as validity (Ainsworth et al., 1993). Since physical fitness is gender-specific, it is important to know the validity of each question for different genders. In all these studies, gender-specific validity was not reported, thus its application in different genders was unknown.

The TOQ took several physiological measures, including maximal oxygen consumption, percent body fat, forced expiratory volume (forced expiratory volume in 1 second/height in meters squared), and accelerometer scores (kcal/day), as physical activity validation criteria. However, the concurrent validity coefficients were low and statistically non-significant ($r \leq .17$, $P > .05$). The correlation of the TOQ with other established PA questionnaires varied between .02 and .92 (Ainsworth et al., 1993).

It is worthwhile to note that the validity coefficients of SPFQ in the present study were very similar to Abadie's Perceived Physical Fitness Scale, as shown in Table 16. Although Abadie's questionnaire was validated between two age groups while the SPFQ of the present study was validated according to gender, the validity coefficients of body composition from both questionnaires were $\alpha \geq -.57$. For the components of muscular strength and flexibility, the validity coefficients yielded by both genders from the SPFQ were also close to the coefficients yielded for the younger sample of Abadie's study.

The similarity between SPFQ and Abadie's Perceived Physical Fitness Scale may partly be due to the same aspects they measure. Both questionnaires were developed to assess health-related physical fitness, with several components in common: flexibility, body composition, and muscular strength. In addition, the ages of the adult participants in these two studies were similar. To be specific, the mean age of the participants from the SPFQ was 47 years old, and that accounted for 63% of the SPFQ participants who matched the younger age group (< 50 years old) of Abadie's study.

It is interesting to note that in Godin's study, fair and fairly low correlations were found in VO_2max consumption and in percent body fat, respectively, with strenuous exercises. Like the SPFQ, the GLTEQ also failed to present a better correlation between the CR component and criterion tests.

In conclusion, due to the poor criterion-related validity of CR items, no question item concerning CR was included in the recommended SPFQ. Eventually, provided the fair to moderate validity coefficients in the other four components, the recommended SPFQ and the associated test-retest reliability and internal consistency are provided in Appendixes G (women) and H (men) for reference.

Table 16

Comparison of criterion-related validity in self-reported physical fitness/ physical activity questionnaires.

Studies	Summary results		
	Variables	Women	Men
SPFQ (Present study)	BC	-.60**	-.57**
	MS	.30**	.31**
	ME	.20	.25*
	FLEX	.39**	.57**
	Variables	Older sample	Younger sample
Perceived Physical Fitness Scale (Abadie, 1988)	CR endurance	.43**	.61**
	Muscular strength	.07*	.47**
	Flexibility	.12*	.53**
	Body composition	-.67**	-.68**

Table 16 (Continued)

		VO_2max	BF
GLTEQ	Light	.04	.06
(Godin & Shephard,	Moderate	.03	.08
1985)	Strenuous	.38*	.21*
	Total	.24*	.13*
	Sweat	.26*	.21*
	Variables	Older sample	Younger sample
TOQ	Questionnaires		
(Ainsworth et al., 1993)	Occupational PA record	.92*	
	7-Day recall	.38*	
	Minnesota Heart Health	.34*	
	CARDIA Occupational	.02	
	Physiological measures	≤ .17	

* $p < .05$; ** $p < .01$

Note. SPFQ = Self-reported Physical Fitness Questionnaire; GLTEQ = Godin Leisure-Time Exercise Questionnaire; TOQ = Tecumseh Occupational Physical Activity Questionnaire.

Limitations

Results from the study support the use of a SPFQ for evaluating four health-related physical fitness components among adults. The present study has several limitations that should be noted:

1. This study was limited to the sample of Hong Kong Chinese adults who reside in the Hong Kong territory, and thus, questions were designed according to the living experience and daily activities in Hong Kong. The transportation system and shopping behaviors are different from other Asian cities. For example, people in Hong Kong usually carry all shopped goods back home by themselves, while people in other cities may merely carry the goods back to their cars and then drive home.
2. The purpose of the questionnaire was to allow adults to estimate their physical fitness levels by self-reporting their performance in functional daily activities. However, some questions were unavoidably querying subjectively (perception) instead of relatively objectively (observation). Hence, the subjectivity of questions should be noted in future studies of questionnaire development.
3. The sample size of N = 169 was not sufficiently large to produce construct validity. Therefore, the present study serves as a pilot study and as a premier reference for future studies.
4. Because educational background, experience in lifestyle and occupational activities, and economic status of the participants were not investigated prior to the study, the extent to which these factors could have affected their responses in SPFQ was unknown.
5. Criterion tests used in the present study may not be the best choice. This may have hindered the criterion-related validity of the question items.
6. Scoring for each component has not yet been investigated. Therefore, norms

cannot be produced.

Recommendations for further study

A validated self-reported questionnaire which could measure or estimate one's health-related physical fitness would be invaluable. Only a questionnaire and a pen are needed to measure one's health-related physical fitness. It is therefore much more convenient than taking part in various laboratory or field tests which are time, labor, and cost demanding. However, due to the low criterion-related validity in CR component in the current study, the SPFQ was found useful to estimate BF, FLEX, MS, and ME components. This caused an insufficiency in developing an all-round self-reported questionnaire for health-related physical fitness. Thus, future studies are necessary to develop questions for evaluating cardiorespiratory fitness.

Furthermore, in future studies, the use of maximal test to measure VO_2max is recommended for a more accurate result. The low validity of CR component in the present study may be affected by using a less accurate VO_2max estimation from a submaximal treadmill test.

Although these preliminary findings support the use of SPFQ to assess people's physical fitness, it still needs to include a larger sample size to examine the construct validity in future studies and to develop norms for each physical fitness component. When norms are developed, people could know their fitness levels compared with others of the same sex and similar age.

Conclusions

The primary aim of developing the SPFQ was to assess health-related physical fitness by self-reporting daily functional ability. A total of 169 Hong Kong Chinese adults (83 women and 86 men) participated in a series of laboratory fitness tests and completed the SPFQ questionnaire twice. Data collected was used for reliability and validity analyses.

Statistical analysis showed very high test-retest reliability analysis and internal consistency for both genders. Criterion-related validity, however, showed an unfavorable result. Except for the question items of the BC component, which yielded moderate correlation with the validation criterion, all other components had low correlations; it was especially low in the CR component. Thus, the investigator attempted to remove items with very low correlations with the criterion before reexamining the criterion-related validity of each item. Since there was no improvement in CR items, all CR items were removed from the recommended SPFQ. Finally, the final version of the recommended SPFQ only included question items concerning the fitness components of body composition, flexibility, muscular strength and muscular endurance, in which 15 items were included in the women's version and nine items in the men's version. Yet, the validity coefficients were overall too low for the questionnaire to be acceptable. In the scientific world, many non-significant statistical findings have significant theoretical and practical values. Therefore, further studies are needed to develop a self-reported fitness questionnaire with satisfactory construct validity that could assess all five components of health-related fitness.

REFERENCES

1981 Canada Fitness Survey [PDF file]. Ottawa, ONT: Canadian Fitness and Lifestyle Research Institute. Available from Canadian Fitness and Lifestyle Research Institute Web site, http://www.cflri.ca

AAHPER. (1965). *AAHPER Youth Fitness Test manual* (Rev. ed.). Washington DC: American Association for Health, Physical Education, and Recreation.

Abadie, B. R. (1988). Construction and validation of a perceived physical fitness scale. *Perceptual and Motor Skills, 67,* 887-892.

ACSM. (2000). *ACSM's guidelines for exercise testing and prescription* (6th ed.). Philadelphia: Lippincott Williams & Wilkins.

ACSM. (2006). *ACSM's guidelines for exercise testing and prescription* (7th ed.). Philadelphia: Lippincott Williams & Wilkins.

Adam, C. V., Klissouras, M., Ravazollo, R., Renson, W., & Tuxworth, W. (1988). Eurofit: *European test of physical fitness*. Rome: Council of Europe, Committee for the Development of Sport.

Ainsworth, B. E., Jacobs, D. R., Leon, A. S., Richardson, M. T., Montoye, H. J. (1993). Assessment of the accuracy of physical activity questionnaire occupational data. *Journal of Occupational Medicine, 35,* 1017-1027.

Anderson, C., Laubscher, S., & Burns, R. (1996). Validation of the short form 36 (sf-36): Health survey questionnaire among stroke patients. *Stroke, 27,* 1812-1816.

Åstrand, P.-O. & Ryhming, I. (1954). A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work. *Journal of Applied Physiology, 7,* 218-221.

Baecke, J. A. H., Burema, H. J., Frijters, J. E. R. (1982). A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *American Journal of Clinical Nutrition, 36,* 936-942.

Ball, S. D. (2005). Interdevice variability in percent fat estimates using the bod pod. *European Journal of Clinical Nutrition, 59*, 996-1001

Ball, S. D., & Altena, T. S. (2004). Comparison of the bod pod and dual energy x-ray absorptiometry in men. *Physiological measurement, 25*, 671-678.

Baumgartnet, T. A., & Jackson, A. S. (1999). *Measurement for evaluation in physical education and exercise science.* (6^{th} ed.). Boston: WCB/McGraw-Hill.

Biaggi, R. R., Wollman, M. W., Niew, M. A., Brener, C. E., Flakoll, P. J., Levenhagen, D. K., et al. (1999). Comparison of air-displacement plethysmography with hydrostatic weighing and bioelectrical impedance analysis for the assessment of body composition in healthy adults. *American Journal of Clinical Nutrition, 69*, 898-903.

Bjorner, J. B., & Kristensen, T. S. (1999). Multi-item scales for measuring global self-rated health. *Research on Aging, 21*, 417-439.

Blair, S. N. (1993). 1993 C. H. McCloy research lecture: Physical activity, physical fitness, and health. *Research Quarterly for Exercise and Sport, 64*, 365-376.

Blair, S. N., Cheng, Y., & Holder, J. S. (2001). Is physical activity or physical fitness more important in defining health benefits? *Medicine and Science in Sports and Exercise, 33*, S379-S399.

Blair, S. N., Kampert, J. B., Kohl, H. W., Barlow, C. E., Macera, C. A., Paffenbarger, R. S., et al. (1996). Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. *Journal of the American Medical Association, 276*, 205-210.

Blair, S. N., Kohl, H. W., Paffenbarger, R. S., Clark, D. G., Cooper, K. H., & Gibbons,

L. W. (1989). Physical fitness and all-cause mortality: A prospective study of healthy men and women. *Journal of American Medical Association, 262*, 2395-2401.

Blair, S. N., Kohl, H. W. I., Barlow, C. E., Paffenbarger, R. S. J., Gibbons, L. W., & Macera, C. A. (1995). Changes in physical fitness and all-cause mortality: A prospective study of healthy and unhealthy men. *Journal of the American Medical Association, 273*, 1093-1098.

Blair, S. N., Wei, M., & Lee, C. D. (1998). Cardiorespiratory fitness determined by exercise heart rate as a predictor of mortality in the aerobics center longitudinal study. *Journal of Sports Sciences, 16*, S47-S55.

Bolanowski, M., & Nilsson, B. E. (2001). Assessment of human body composition using dual-energy x-ray absorptiometry and bioelectrical impedance analysis. *Medical Science Monitor, 7*, 1029-1033.

Bosy-Westphal, A., Mast, M., Eichhorn, C., Becker, C., Kutzner, D., Heller, M., et al. (2004). Validation of air-displacement plethysmography for estimation of body fat mass in healthy elderly subjects. *European Journal of Nutrition, 42*, 207-216.

Brock, S. J., & Fittipaldi-Wert, J. (2005). Making the fitness connection. *Teaching Elementary Physical Education, 16,* 15-16.

Bruce, R. A., Kusumi, F., & Hosmer, D. (1973). Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *American Heart Journal, 85*, 546-562.

Canadian Fitness and Lifestyle Research Institute. (n.d.). *1981 Canadian fitness survey.* Retrieved December 19, 2005, from http://www.cflri.ca

Canadian Society for Exercise Physiology. (1998). *The Canadian physical activity,* *fitness & lifestyle appraisal: CSEP's plan for healthy active living* (2^{nd} ed.). Ottawa: Author.

Centers for Disease Control and Prevention. (n.d.). *CDC scientific data index.* Retrieved December 14, 2005, from http://wonder.cdc.gov

Clasey, J. L., & Gater, D. R. (2005). A comparison of hydrostatic weighing and air displacement. *Archives of Physical Medicine and Rehabilitation, 86*, 2106-2113.

Cooper, K. H. (1968). A means of assessing maximal oxygen intake. *Journal of the American Medical Association, 203*, 201-204.

Cooper Institute, The. (n.d.). *FITNESSGRAM.* Retreived January 2, 2006, from http://www.cooperinstitute.org/products/grams/index.cfm

Council of Europe. *EUROFIT.* (n.d.) Retrieved December 13, 2005, from http://www.coe.int

Cureton, K. J., Sloniger, M. A., O'Bannon, J. P., Black, D. M., & McCormack, W. P. (1995). A generalized equation for prediction of $VO2_{peak}$ from 1-mile run/ walk performance. *Medicine and Science in Sports and Exercise, 27*, 445-451.

Demerath, E. W., Guo, S. S., Chumlea, W. C., Towne, B., Roche, A. F., & Siervogel, R. M. (2002). Comparison of percent body fat estimates using air displacement plethysmography and hydrodensitometry in adults and children. *Journal of Internaional Association for the Study of Obesity, 26*, 389-397.

DiPietro, L., Kohl, H. W., Barlow, C. E., & Blair, S. N. (1998). Improvements in cardiorespiratory fitness attenuate age-related weight gain in healthy men and women: The aerobics center longitudinal study. *International Journal of Obesity, 22*, 55-62.

Doyne, E. J., Ossip-Klein, D. J., Bowman, E. D., Osborn, K., McDougall-Wilson, I. B., & Neimeyer, R. A. (1987). Running versus weight lifting in the treatment of depression. *Journal of Consulting & Clinical Psychology, 55*, 748-754.

Dunin, J. V. G. A., & Rahaman, M. M. (1967). The assessment of the amount of fat in the human body from measurements of skinfold thickness. *British Journal of Nutrition, 21*, 681-689.

Ekelund, L. G., Haskell, W. L., Johnson, J. L., Whaly, F. S., Criqui, M. H., & Sheps, D. S. (1988). Physical fitness as a predictor of cardiovascular mortality in asymptomatic North American men: The Lipid Research Clinics mortality follow-up study. *New England Journal of Medicine*, 319, 1379-1384.

Erikssen, J. (1986). Physical fitness and coronary heart disease morbidity and mortality: A prospective study in apparently healthy, middle-aged men. *Acta Medica Scandinavica Supplementum*, 711, 189-192.

Federal Prevention Activities. (1994, March). *Federal prevention activities*. Retrieved January 10, 2006, from http://www.crisny.org/health/us/health6.html

Fitnessgram. (2003). Retrieved January 2, 2006, from http://www.fitnessgram.net.

FitzGerald, S. J., Barlow, C. E., Kampert, J. B., Morrow, J. R., Jackson, A. W., & Blair, S. N. (2004). Muscular fitness and all-cause mortality: Prospective observations. *Journal of Physical Activity and Health, 1,* 7-18.

Flakoll, P. J., Kent, P., Neyra, R., Levenhagen, D., Chen, K. Y., & Ikizler, T. A. (2004). Bioelectrical impedance vs air displacement plethysmography and dual-energy x-ray absorptiometry to determine body composition in patients with end-stage renal disease. *Journal of Parenteral and Enteral Nutrition, 28*, 13-21.

Frisard, M. I., Greenway, F. L., & Delany, J. P. (2005). Comparison of methods to assess body composition changes during a period of weight loss. *Obesity Research, 13*, 845-854.

Gama, E. V., Damian, J., Molino, J. P. d., Lopez, M. R., Perez, M. L., & Iglesias, F. J. G. (2000). Association of individual activities of daily living with self-rated health in older people. *Age and Ageing, 29*, 267.

George, J. D., Stone, W. J., & Burkett, L. N. (1997). Non-exercise VO2max estimation for physically active college students. *Medicine and Science in Sports and Exercise, 29*, 415-423.

George, J. D., Vehrs, P. R., Allsen, P. E., Fellingham, G. W., & Fisher, A. G. (1993). Vo2max estimation from a submaximal 1-mile track jog for fit college-age individuals. *Medicine and Science in Sports and Exercise, 25*, 401-406.

Glickman S. G., Marn, C. S., Supiano, M. A., & Dengel, D. R. (2004). Validity and reliability of dual-energy x-ray absorptiometry for the assessment of abdominal adiposity. *Journal of Applied Physiology, 97*, 509-514.

Godin, G., & Shephard , R. J. (1985). A simple method to assess exercise behavior in the community. *Canadian Journal of Applied Sport Science, 10,* 141-146.

Going, S. B., Massett, M. P., Hall, M. C., Bare, L. A., Root, P. A., Williams, D. P., et al. (1993). Detection of small changes in body composition by dual-energy x-ray absorptiometry. *American Journal of Clinical Nutrition, 57*, 845-850.

Golding, L. (Ed.). (2000). *YMCA fitness testing and assessment manual.* Champaign, Il: Human Kinetics.

Golding, L., Myers, C., & Sinning, W. (1989). *Y's way to physical fitness.* Champaign, IL: Human Kinetics.

Hansen, R. D., Raja, C., Aslani, A., Smith, R. C., & Allen, B. J. (1999).

Determination of skeletal muscle and fat-free mass by nuclear and dual-energy x-ray absorptiometry methods in men and women aged 51-84y. *American Journal of Clinical Nutrition, 70*, 228-233.

He, M., Tan, K. C., Li, E. T., & Kung, A. W. (2001). Body fat determination by dual energy x-ray absorptiometry and its relation to body mass index and waist circumference in Hong Kong Chinese. *Journal of the International Association for the Study of Obesity, 25*, 748-752.

Heberlein, T. A., & Baumgartner, R. (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. *American Sociological Review, 43*, 447-462.

Heyward, V. H. (2002). *Advanced fitness assessment and exercise prescription* (4 ed.). Champaign, IL: Human Kinetics.

Hoeger, W. W. K., & Hopkins, D. R. (1992). A comparison of the sit and reach and the modified sit and reach in the measurement of flexibility in women. *Research Quarterly for Exercise and Sport, 63*, 191-195.

Hong Kong Department of Health. *Vital statistics 2004*. Retrieved December 19, 2005, from http://www.info.gov.hk/dh/diseases/vitstat.htm

Hootman, J. M., Macera, C. A., Ainsworth, B. E., Martin, M., Addy, C. L., & Blair, S. N. (2001). Association among physical activity level, cardiorespiratory fitness, and risk of musculoskeletal injury. *American Journal of Epidemiology, 154*, 251-258.

Houtkooper, L. B., Going, S. B., Sproul, J., Blew, R. M., & Lohman, T. G. (2000). Comparison of methods for assessing body-composition changes over 1 y in postmenopausal women. *American Journal of Clinical Nutrition, 72*, 401-406.

Hui, S. & Morrow, J. R. (2001). Level of participation and knowledge of physical activity in Hong Kong Chinese adults and their association with age. *Journal of Aging and Physical Activity, 9*, 372-385.

Hui, S. & Yuen, P. Y. (2000). Validity of the modified back-saver sit-and-reach test: A comparison with other protocols. *Medicine and Science in Sports and Exercise, 32*, 1655-1659.

Jackson, A. S., Blair, S. N., Mahar, M. T., Wier, L. T., Ross, R. M., & Stuteville, J. E. (1990). Prediction of functional aerobic capacity without exercise testing. *Medicine and Science in Sports and Exercise, 22*, 863-870.

Jackson, S., Owen, P., Bender, B., Senner, J., Davis, B., Leff, M., et al. (1995). Leads from the morbidity and mortality weekly report, Atlanta, GA: Health-related quality-of-life measures-United States, 1993. *Journal of the American Medical Association, 273,* 1084-1085.

Jackson, C. J., & Furnham, A. (2000). *Designing and analysing questionnaires and surveys*. London: Athenaeum Press.

Jacobs, D. R. Jr., Ainsworth, B. E., Hartman, T. J., Leon, A. S. (1993). A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Medicine and Science in Sports and Exercise, 25,* 81-91.

Japan Ministry of Education, Culture, Sports, Science and Technology. (2005). *A mind and body sustaining a "zest for living"*. Tokyo: National Printing Bureau.

Johansson, A. G, Forslund, A., Sjodin, A., Hallmin, H., Hambraeus, L., & Ljunghall, S. (1993). Determination of body composition: A comparison of dual-energy x-ray absorptiometry and hydrodensitometry. *American Journal of Clinical Nutrition, 57*, 323-326.

Jwo, C., Kirsten, W. (2000). Country Profile: Health promotion in Taiwan: Leadership from government and academia. *Journal of American of Health Promotion, 3*. Retrieved January 10, 2006, from http://www.healthpromotionjournal.com

Kalton, G., Roberts, J., & Holt, D. (1980). The effects of offering a middle response option with opinion questions. *The Statistician, 29*, 65-78.

Katzmarzyk, P. T., Church, T. S., & Blair, S. N. (2004). Cardiorespiratory fitness attenuates the effects of the metabolic syndrome on all-cause and cardiovascular disease mortality in men. *Archives of Internal Medicine, 164*, 1092-1097.

Kennedy, J. F. (1960). The soft American. *Sports Illustrated,* 15-23.

Kitano, T., Kitano, N., Inomoto, T., & Futatsuka, M. (2001). Evaluation of body composition using dual-energy x-ray absorptiometry, skinfold thickness and bioelectrical impedance analysis in Japanese female college students. *Journal of Nutritional Science and Vitaminology, 47*, 122-125.

Kline, G. M., Porcari, J. P., Hintermeister, R., & Freedson, J. M. (1987). Estimation of VO_2max from a one-mile track walk, gender, age, and body weight. *Medicine and Science in Sports and Exercise, 19*, 253-259.

Koda, M., Ando, F., Niino, N., Tsuzuku, S., & Shimokata, H. (2000). Comparison between the air displacement method and dual energy x-ray absorptiometry for estimation of body fat. *Journal of Epidemiology, 10*, S82-S89.

Kohrt, W. M. (1995). Body composition by DXA: Tried and true? *Medicine and Science in Sports and Exercise, 27*, 1349-1353.

Kolkhorst, F. W., & Dolgener, F. A. (1994). Nonexercise model fails to predict aerobic capacity in college students with high VO_2peak. *Research Quarterly* *for Exercise and Sport, 65*, 78-83.

Lawlor, D. A., Bedford, C., Taylor, M., & Ebrahim, S. (2002). Agreement between measured and self-reported weight in older women. Results from the British women's heart and health study. *Age and Ageing, 31*, 169.

Levenhagen, D. K., Borel, M. J., Welch, D. C., Piasecki, J. H., Piasecki, D. P., Chen, K. Y., et al. (1999). A comparison of air displacement plethysmography with three other techniques to determine body fat in healthy adults. *Journal of Parenteral and Enteral Nutrition, 23*, 293-299.

Maddaalozzo, G. F., Cardinal, B. J., & Snow, C. A. (2002). Concurrent vaoidity of the bod pod and dual energy x-ray absorptiometry techniques for assessing body composition in young women. *Journal of the American Dietetic Association, 102*, 1677-1679.

Mahar, M. T., Rowe, D. A., Parker, C. R., Mahar, F. J., Dawson, D. M., & Holt, J. E. (1997). Criterion-referenced and norm-referenced agreement between the mile run/ walk and pacer. *Measurement in physical education and exercise science, 1*, 248-258.

Marsh, H. W., & Redmayne, R. S. (1994). A multidimensional physical self-concept and its relations to multiple components of physical fitness. *Journal of Sport & Exercise Psychology, 16*, 43-55.

Marsh, H. M., Richards, G. E., Johnson, S., Roche, L., & Tremayne, P. (1994). Physical self-description questionnaire: Psychometric properties and a multitrait-multimethod analysis of relations to existing instruments. *Journal of Sport & Exercise Psychology, 16,* 270-305.

Marsh, H. M. (1996). Physical self description Questionnaire: Stability and discriminant validity. *Research Quarterly for Exercise and Sport, 67*, 249-64.

McCrory, M. A., Gomez, T. D., Bernauer, E. M., & Mole, P. A. (1995). Evaluation of a new air displacement plethysmograph for measuring human body composition. *Medicine and Science in Sports and Exercise, 27*, 1686-1691.

Ministry of Education, Culture, Sports, Science and Technology, Japan. (n.d.). *「平成１６年度体力・運動能力調査」の概要* [Summary of the Physical strength and motor fitness survey]. Retrieved December 19, 2005, from http://www.mext.go.jp/b_menu/houdou/sports.htm

Morrow, J. R., Jackson, A. W., Disch, J. G., & Mood, D. P. (2000). *Measurement and evaluation in human performance* (2nd ed.). Champaign, IL: Human Kinetics.

Nieman, D. C. (1998). Physical activity components of health. In H. Gilly & K. Reel (Eds.), *The exercise-health connection* (pp. 3-19). Champaign, IL: Human Kinetics.

Norcross, J., & Van Loan, M. D. (2004). Validation of fan beam dual energy x ray absorptiometry for body composition assessment in adults aged 18-45 years. *British Journal of Sports Medicine, 38*, 472-476.

Panotopoulos, G., Ruiz, J. C., Guy-Grand, B., & Basdevant, A. (2001). Dual x-ray absorptiometry, bioelectrical impedance, and near infrared interactance in obese women. *Medicine and Science in Sports and Exercise, 33*, 665-670.

Pate, R., Ross, J., Dotson, C., & Gilbert, G. (1985). The new forms: A comparison with the 1980 AAHPERD norms. *Journal of physical education, recreation, and dances, 56*, 70-72.

Peart, N. D., Marsh, H. W., Richards, G. E. (n.d.) *PSDQ - Physical self description questionnaire.* Retrieved March 21, 2007, from http://self.uws.edu.au/Instruments/PSDQ%20PACKAGE/PSDQ.htm

Radley, D., Gately, P. J., Cooke, C. B., Carroll, S., Oldroyd, B., & Truscott, J. G. (2003). Estimates of percentage body fat in young adolescents: A comparison of dual-energy x-ray absorptiometry and air displacement. *European Journal of Clinical Nutrition, 57*, 1402-1410.

Rosandich, T. (1999, Winter). International physical fitness test. *The Sport Journal, 2*(1). Retrieved January 6, 2006, from http://www.thesportjournal.org/1999Journal/Vol2-No1/fitness.asp

Rosandich, T. (2003, Winter). The international physical fitness test manual. *The Sport Journal, 6*(1). Retrieved January 6, 2006, from http://www.thesportjournal.org/2003Journal/Vol6-No1/IPFT_Manual.asp

Safrit, M. J., & Wood, T. M. (1995). *Introduction to measurement in physical education and exercise science* (3rd ed.). Boston, Mass: WCB/McGraw-Hill.

Sallis, J. F., Haskell, W., Wood, P. (1985). Physical activity assessment methodology in the Five-City Project. *American Journal of Epidemiology, 121,* 91-106.

Sardinha, L. B., Lohman, T. G., Teixeira, P. J., Guedes, D. P., & Going, S. B. (1998). Comparison of air displacement plethysmography with dual-energy x-ray absorptiometry and 3 field methods for estimating body composition in middle-aged men. *American Journal of Clinical Nutrition, 68*, 786-793.

Shephard, R. J. (2003). Limits to the measurement of habitual physical activity by questionnaires. *British Journal of Sports Medicine, 37*, 197-206.

Shibuya, M., Kamizato, M., & Tachi, K. (2005). 健康指標としてのウエスト値とその関連要因について [Waist circumference as a health index and its related factors]. *Sanyo Eiseigaku Zasshi, 47*, 122-127.

Siegle, D. (n.d.). *Likert scale*. Retrieved July 5, 2005, from www.gifted.uconn.edu/siegle/research/instrument%20reliability%20and%20v alidity/likert. html

Slattery, M. L., & Jacobs, D. R. (1988). Physical fitness and cardiovascular disease mortality: The U. S. Railroad Study. *American Journal of Epidemiology, 127,* 571-580.

Sletto, R. F. (1940). Pretesting of questionnaires. *American Sociological Review, 5,* 193-200.

Soriano, J. M., Ioannidou, E., Wang, J., Thornton, J. C., Horlick, M. N., Gallagher, D., (2004). Pencil-beam VS fan-beam dual-energy x-ray absorptiometry comparisons across four systems: Body composition and bone mineral. *Journal of Clinical Densitometry, 7,* 281-289.

St-Onge, M. P., Wang, Z., Horlick, M., Wang, J., & Heymsfield, S. B. (2004). Dual-energy x-ray absorptiometry lean soft tissue hydration: Independent contributions of intro- and extracellular water. *American Journal of Physiology, endocrinology and metabolism, 287,* E842-E827.

Stuck, A. E., Elkuch, P., Dapp, U., Anders, J., Iliffe, S., & Swift, C. G. (2002). Feasibility and yield of a self-administered questionnaire for health risk appraisal in older people in three European countries. *Age and Ageing, 31,* 463.

Svendsen, O. L., Haarbo, J., Heitmann, B. L., Gotfredsen, A., & Christiansen, C. (1991). Measurement of body fat in elderly subjects by dual-energy x-ray absorptiometry, bioelectrical impedance, and anthropometry. *American Journal of Clinical Nutrition, 53,* 1117-1123.

Taiwan Ministry of Education Physical Fitness Passport. (n.d.). Retrieved January 10, 2006, from http://www.fitness.org.tw

Teo, W. F., & Dale, B. G. (1997). Self-assessment: Methods, management and process. *Journal of Engineering Manufacture, 211*(B5), 365-376.

Tikkanen, H. O., Hamalainen, E., Sarna, S., Adlercreutz, H., & Harkonen, M. (1998). Associations between skeletal muscle properties, physical fitness, physical activity and coronary heart disease risk factors in men. *Atherosclerosis, 137*, 377-389.

Trochim, W. M. K. (2005). *Research methods knowledge base.*

Turner, P. (2000). Types of activity, fitness levels and calcium intake amongst 14-16 year old scholars: Sufficient for bone health? *Advances in Physiotherapy, 2*, 51-62.

United States Department of Health and Human Services. (1996). *Physical Activity and Health: A Report of the Surgeon General.* Atlanta, GA: Author.

VanLoan, M. D., & Mayclin, P. L. (1992). Body composition assessment: Dual-energy x-ray absorptiometry (DEXA) compared to reference methods. *European Journal of Clinical Nutrition, 46*, 125-130.

Vescovi, J. D., Zimmerman, S. L., Miller, W. C., Hildebrandt, L., Hammer, R. L., & Fernhall, B. (2001). Evaluation of the bod pod for estimating percentage body fat in a heterogeneous group of adult humans. *European Journal of Applied Physiology, 85*, 326-332.

Villeneuve, P. J., Morrison, H. I., Craig, C. L., & Schaubel, D. E. (1998). Physical activity, physical fitness, and risk of dying [electronic version]. *Epidemiology, 9*, 626.

Wagner, D. R., Heyward, V. H., & Gibson, A. L. (2000). Validation for air displacement plethysmography for assessing body composition. *Medicine and Science in Sports and Exercise, 32*, 1339-1344.

Wang, Y., Rimm, E. B., Stampfer, M. J., Willett, W. C., & Hu, F. B. (2005).

Comparison of abdominal adiposity and overall obesity in predicting risk of type 2 diabetes among men. *American Journal of Clinical Nutrition, 81*, 555-563.

Warburton, D. E. R., Gledhill, N., & Quinney, A. (2001). Musculoskeletal fitness and health. *Canadian Journal of Applied Physiology, 26*, 217-237.

Wier, L. T., Jackson, A. S., & Pinkerton, M. B. (1989). Evaluation of the NASA/JSC health related fitness program. *Aviation, Space, and Environmental Medicine, 60*, 438-444.

Wildman, R. P., Gu, D., Renolds, K., Duan, X., Wu, X., & He, J. (2005). Are waist circumference and body mass index indendently associated with cardiovascular disease disk in Chinese adults? *American Journal of Clinical Nutrition, 82*, 1195-1202.

Williford, H. N., Scharff-Olson, M., Wang, N., Blessing, D. L., Smith, F. H., & Duey, W. J. (1996). Cross-validation of non-exercise predictions of VO2peak in women. *Medicine and Science in Sports and Exercise, 28*, 926-930.

World Health Organization. (2000). *The Asia-Pacific perspective: Re-defining obesity and its treatment.* Australia: International Diabetes Institute, WHO Regional Office for the Western Pacific.

Wuest, D. A., & Bucher, C. A. (2003). *Foundations of physical education, exercise science, and sport* (14th ed.). New York: McGraw-Hall.

Yang, Y., Zhu, W. D., & Paton, N. I. (2004). Comparison of dual-energy x-ray absorptiometry machines for measuring fat distribution changes of hiv-associated lipodystrophy. *Antiviral Therapy, 9*, 771-778.

Zhu, S., Heymsfield, S. B., Toyoshima, H., Wang, Z., Pietrobelli, A., & Heshka, S. (2005). Race-ethnicity-specific waist circumference cutoffs for identifying cardiovascular disease risk factors. *American Journal of Clinical Nutrition, 81*, 409-415.

許世全，(2002)。**體適能評估理論及應用**。香港：中國香港體適能總會。

國家體育總局科教司 [China National Sports Council]，(2000)。**中國國民體質監測系統的研究**。

Appendix A

體適能自我評估
Self-Reported Physical Fitness Questionnaire (SPFQ)

背景和目的：

直到目前爲止，要有效地評估一個人的體適能情況（physical fitness），必須透過各種實驗室測試（laboratory tests 或 field tests）才能知道結果。雖然這是目前較準確的做法，但卻需要很多人力、物力和時間，侷限了這方面的普及性。爲了讓更多人知道自己的體適能情況，我們希望能設計一份能有效地評估體適能的問卷，讓大家在不用到指定的場地、不需要專業人士的指導、和沒有成本高昂的科學儀器和器材的環境下，也能得知自己的體適能情況。

此問卷專爲年齡介乎三十至六十歲的中年男女設計。

問卷內容：

此問卷將會評估以下量度體適能的五個項目：

1. <u>心肺耐力 (Cardiorespiratory Endurance)</u>
   - 心、肺以及循環系統提供身體所需氧氣給肌肉的有效程度
2. <u>肌力 (Muscular Strength)</u>
   - 由肌肉或肌肉群組在一次肌肉收縮中所能發出的最大力量
3. <u>肌耐力 (Muscular Endurance)</u>
   - 在一指定時間內，肌肉能維持次極限力度(submaximal force levels))的能力，而沒有明顯的關節活動
4. <u>柔軟度 (Flexibility)</u>
   - 在不引致受傷的情況下，順暢地活動關節的能力
5. <u>脂肪比例 (Body Composition)</u>
   - 對身體中的脂肪比例（百分比）的評估。

答卷方法：

1. 此問卷共有兩部份。第一部份爲體適能的自我評估；第二部份爲答卷者的基本資料。
2. 請細閱題目的日常處景對自己表現和能力的描述，再以「✓」號選出適合的<u>真確程度</u>。
3. 要評估自己的體適能情況，請務必回答所有問題。
4. 在回答各問題時，請考慮剛過去三個月的情況。
5. 有些題目比較相似，故此請小心閱讀每條問題。
6. 請必須及只能選取<u>一個</u>答案。
7. 請於完卷後從頭檢視各題，以確保沒有問題遺漏。
8. 這份問卷是不記名的。所有問卷內的資料均只作學術研究之用，絕對保密，並會於研究完成後銷毀。

時間：
約需 20 分鐘

請注意：

- 此問卷共有五個範疇；問題主要針對日常生活中會遇到的情況，*盡可能*不會問答題者對自己的觀感
- 此問卷中的用詞、文法等可能有不恰當、不清晰和難以理解的地方，請給予建議
- 請於指定空格內寫下改善建議；如在「其他建議」一欄提出建議，請以 A1 表示 A 組的第一條，B3 表示 B 組的第三條，如此類推
- 灰色字體的文字並不會在真正的問卷中列印出來
- 在真正的問卷中，問題會以隨機方式排列，並不一定會按測試組別顯示
- 請給予任何建議或批評，例如問卷的形式、外貌……
- 有 * 號的問題是以相反評分計算（reversed scoring）
- 請於＿＿＿＿＿＿＿＿＿＿前交回
- 如有疑問，請致電 9873-8070 或電郵 specandy@cuhk.edu.hk，與戚小姐（Candy）

一・問卷部份

A. Cardiorespiratory Endurance （心肺耐力）								
	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份針對心肺耐力，要求答卷者對其平日的活動能力和表現作出評分。 除了有*號的問題外，預期心肺耐力較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
1	我不停地步行上三十級樓梯（約一至兩層樓），也不會喘氣。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
2	我不停地步行落三十級樓梯，也不會喘氣。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
3	以極快速跑約 50 米後（如追巴士），我的心跳和呼吸都能於一分鐘內回復正常。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
4	我能以中等並且保持平均的速度跑畢四百米（圍繞一個標準田徑場一周）。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

		II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份針對心肺耐力。要求答卷者對其平日的活動能力和表現作出評分。 除了有*號的問題外，預期心肺耐力較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
5	我在一個十米的距離內急步來回十次後，也能夠於一分鐘內恢復正常呼吸和心跳。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
6	我平日走路的節奏比一般同年紀和同性別的人士快。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
7	*我平日走路因為前面的人走得慢而不能再走快點。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
8	*我步行時被後面的人追上。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
9	*我在平路上急步走一分鐘會喘氣。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
10	*蹲下然後返回站立姿勢時，我會有喘氣的情況。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

		II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份針對心肺耐力。要求答卷者對其平日的活動能力和表現作出評分。 除了有*號的問題外，預期心肺耐力較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
11	*跪在地上然後返回站立姿勢時，我會有喘氣的情況。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
12	*跪在地上然後返回站立姿勢時，我心跳會加速。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
13	*步行時，我會比一般同年紀的人喘氣。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
14	與一般人比較，我較能輕鬆地完成二十分鐘的緩步跑。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
15	與一般人比較，我較能輕鬆地完成二十分鐘的中速跑。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
16	與一般人比較，我能較輕鬆地快跑五步。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

		II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份針對心肺耐力，要求答卷者對其平日的活動能力和表現作出評分。 除了有*號的問題外，預期心肺耐力較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
17	與一般人比較，我快跑五步後不會喘氣。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

意見 / 建議：

B. Muscular Strength （肌力量）								
	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者對自己日常生活中會用到肌肉力量的動作評分。 除了有*號的問題外，預期肌力量較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
1	我起床時，即使不用手撐著床或扶著任何物件，也能自己坐起來。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
2	我蹲在地上，即使完全不用扶著任何人或物件也能輕易地站起來。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
3	我坐在沙發上，雙腳平放在地，不用雙手的協助也能輕易地站起來。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
4	我盤膝坐在地上後，*只用雙腿*的力量便能輕易地站起來。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
5	我能輕易地登上公共交通工具（如巴士、小巴、電車等），而不需身體其他部分和他人的協助。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者對自己日常生活中會用到肌肉力量的動作評分。 除了有*號的問題外，預期肌力量較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
6	我不停地向上行三十級樓梯（例如從地下鐵或火車的月台步行上大堂），並不會感到吃力。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
7	我能夠輕易地用單手提起一張椅子。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
8	我能夠輕易地以單手拿起一袋五公斤的米。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
9	我能夠輕易地以單手拿起一袋八公斤的米。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
10	我能夠輕易地以單手拿起一罐 2.5 公升的罐裝食油。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
11	我能夠輕易地以單手拿起一排 8 罐裝的 355 毫升的汽水或啤酒（共 2.8 公升）。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者對自己日常生活中會用到肌肉力量的動作評分。 除了有*號的問題外，預期肌力量較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
12	我能夠輕易地以單手拿起一樽 5 公升的食油。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
13	我能夠輕易地以單手拿起一盒 2.5 公斤的洗衣粉。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
14	我能夠輕易地以單手拿起一瓶 6 公升的蒸餾水或礦泉水。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
15	我能夠輕易地以單手拿起一包六罐裝（355 毫升 x6，共 2.13 公升）的汽水或啤酒。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
16	我能夠輕易地以單手拿起一排六包裝（250 毫升 x6，共 1.5 公升）的紙包飲品。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
17	我能輕易地用雙手搬起一張單座位梳化。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者對自己日常生活中會用到肌肉的動作評分。 除了有*號的問題外，預期肌力量較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
18	*對我來說，用雙手拿起一個普通大小的圓形西瓜是一件很吃力的事。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

意見 / 建議：

C. Muscular Endurance （肌耐力）								
	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者對自己日常生活中會較長時間用到肌肉的動作評分。 除了有*號的問題外，預期肌耐力較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
1	我不停地步行上三十級樓梯（例如：從地下鐵或火車的月台步行往大堂），並不會感到疲累。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
2	我不停地步行落三十級樓梯（例如：從地下鐵或火車的月台步行往大堂），並不會感到疲累。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
3	平日購物後（包括超級市場），我能夠用同一隻手提著貨品步行超過一分鐘。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
問 4 至 12：我不能夠輕易地以同一隻手拿著……								
4	*一包五公斤的米超過一分鐘。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
5	一包八公斤的米超過一分鐘。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

	I. 問題	II. 自我評分					III. 相關性(請圈出選擇)	IV. 建議
	此部份要求答卷者對自己日常生活中會較長時間用到肌肉的動作評分。 除了有*號的問題外，預期肌耐力較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
6	*我不能夠以同一隻手拿著一罐 2.5 公升的罐裝食油超過一分鐘。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
問 4 至 12：我不能夠輕易地以同一隻手拿著……								
7	*我不能夠以同一隻手拿著一排 3 x 900 公升（共 2.7 公升）的食油超過一分鐘。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
8	*我不能夠以同一隻手拿著一樽 5 公升的食油超過一分鐘。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
9	*一盒 2.5 公斤的洗衣粉超過一分鐘。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
10	一瓶 6 公升的蒸餾水或礦泉水超過一分鐘。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
11	*一包六罐裝（355 毫升 x 6，共 2.13 公升）的汽水或啤酒超過一分鐘。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者對自己日常生活中會用到肌肉的動作評分。 除了有*號的問題外，預期肌力量較弃的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
12	*一排六包裝（250 毫升 x 6，共 1.5 公升）的紙包飲品超過一分鐘。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
問 13 至 16：我能夠輕易地……								
13	以單手搬動一張椅子，並移動最少兩米距離。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
14	以雙手搬起一張單人梳化，並移動兩米距離。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
15	單手提著一部手提電腦超過三分鐘而不會感到疲累（以正常的重量計算）。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
16	*於放下重物後（如一包五公斤的米），我便很難再拿起。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
意見 / 建議：								

D. Flexibility （柔軟度）								
	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者給自己評估一下在日常生活中做到題目所指定的動作的能力。 除了有*號的問題外，預期柔軟度較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
問 1 至 3：在沒有任何人或物件的協助下……								
1	我能夠自行背上背包而不會感到困難。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
2	我能輕易地先把背包的一條肩帶放在其中一肩膀上，再把另一條肩帶放在另一肩膀上。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
3	我能輕易地把背包的兩條肩帶以任何方式放上兩肩。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
4	背部兩肩胛中間發癢時，我能舉起右手抓癢。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
5	背部兩肩胛中間發癢時，我能舉起左手抓癢。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者給自己評估一下在日常生活中做到題目所指定的動作的能力。 除了有*號的問題外，預期柔軟度較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
6	後腰中間發癢時，我能用右手從下而上來抓癢。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
7	後腰中間發癢時，我能用左手從下而上來抓癢。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
8	我能輕易以一步跨上三級樓梯（即從第一級跨上第三級）。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
9	我坐在椅子上，能夠抱著腳來剪腳甲。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
10	我在洗澡時能用右手洗刷腰背所有位置。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
11	我在洗澡時能用左手洗刷腰背所有位置。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者給自己評估一下在日常生活中做到題目所指定的動作的能力。 除了有*號的問題外，預期柔軟度較差的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
12	在巴士、火車或地鐵車廂中，我能舉手並能完全地緊握到車廂天花的扶手。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
13	在超級市場中，我能輕易伸直手拿取比自己高約 1.5 尺（45 厘米）的物件。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
14	收拾床舖時，我做出揚起被單的動作沒有感到困難。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
15	坐在椅子上，雙腳放在地上，我能輕易地向前彎腰繫鞋帶。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
16	*搬動沉重的物件時，我感到很容易拉傷肌肉或關節。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
17	我能以站立姿勢，彎腰拾起地上不太重的物件而不用屈膝。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

意見 / 建議：

E. Body Composition （脂肪比例）

	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者給自己評估一下在日常生活中做到題目所指定的動作的能力。 除了有*號的問題外，預期肥胖的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
1	坐在椅上時，挺直腰背、上身與大腿成九十度角的動作，對我來說完全無問題。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
2	坐在椅上挺直上身時，腹部完全不會接觸到大腿。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
3	我能夠在巴士車廂中的樓梯輕鬆地上落。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
4	*試穿長褲子時，我的大腿位置全塞不下。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
5	我蹲下繫鞋帶時完全沒有困難。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
6	我盤膝坐地完全沒有困難。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者給自己評估，下在日常生活中做到題目所指定的動作的能力。 除了有*號的問題外，預期肥胖的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
7	坐公共交通工具時，我只需一個座位已經十分足夠了。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
8	*在公共交通工中，一個座位並不足夠我坐。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
9	我保持上身挺直地站立時，低下頭能望到自己的腳趾。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
10	我的動作敏捷。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
11	*我活動時感到身形影響到我的活動能力和敏捷性。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
12	*我原地跳時會感到自己很重。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

	I. 問題	II. 自我評分					III. 相關性（請圈出選擇）	IV. 建議
	此部份要求答卷者給自己評估一下在日常生活中做到題目所指定的動作的能力。 除了有*號的問題外，預期肥胖的答卷者會選擇「從不真確」。 「經常真確」表示問題能完全準確地描述自己的情況；「從不真確」表示問題沒有一點對自己正確的描述。	經常真確	通常真確	有時真確	很少真確	從不真確		
13	*我的腰圍比一般同身高的人士爲大。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
14	*我的下圍比一般同身高的人士爲大。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
15	*兩腿直立時，兩腳掌不能緊貼。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	
16	*與身高相若並且同性別的人士比較，我通常比他/她們穿較大號的衣服。						A. 是，不需修改 B. 是，建議修改（如欄 IV） C. 否	

意見 / 建議：

其他意見 / 建議：

Appendix B

體適能自我評估
Self-Reported Physical Fitness Questionnaire (SPFQ)

爲什麼有這份問卷？

我們是香港中文大學體育運動科學系的研究員，現正協助進行一次以問卷方式調查香港成年人體適能狀況（physical fitness）的研究。是項研究結果將有助香港制訂健康與體適能的推廣策略。

很多外國和本地的研究都顯示，健康和體適能有著密切的關係。隨著近年大眾對健康意識的提高，體適能狀況亦越來越受到重視。因此要知道自己哪方面的體能需要改善，最好的途徑是先瞭解自己的體適能狀況。

爲了讓更多人知道自己的體適能狀況，我們現設計了一份評估體適能的問卷，同時亦讓我們瞭解大家對自己體適能狀況的認知。

問題是怎樣的？

此問卷專爲年齡介乎二十四至六十歲的成年人設計。問題主要環繞日常生活中所遇到的有關體能的問題，並不會涉及私人問題。問卷是不記名的。所有問卷內容均只作學術研究之用，絕對保密，並會於研究完成後銷毀。

如何答卷？

1. 請考慮剛過去三個月的身體狀況，並回答所有問題。每題只能選取一個答案。
2. 請細閱題目所預設的的日常處境對自己體能表現和能力的描述，以「✓」號選出適合的真確程度。
3. 有些題目比較相似，故此請小心閱讀每條問題。

時間：

需時約 15-20 分鐘

聯絡：

如對本問卷有任何疑問，請聯絡：
香港中文大學體育運動科學系
戚芷君小姐：2609-8069
許世全教授：2609-6081

＊＊（本頁可於問卷完成後取去）＊＊

香港中文大學體育運動科學系

體適能自我評估問卷

參加者編號：＿＿＿＿＿＿＿＿	填寫日期：＿＿＿年＿＿月＿＿日

此問卷專爲年齡介乎二十四至六十歲的成年人設計。問題主要環繞日常生活中所遇到的有關體能的問題，並不會涉及私人問題。所有問卷內容均只作學術研究之用，絕對保密，並會於研究完成後銷毀。整份問卷需時約 15-20 分鐘。

1. 請考慮剛過去三個月的身體狀況，並回答所有問題。每題只能選取一個答案。
2. 請細閱題目所預設的的日常處境對自己體能表現和能力的描述，以「✓」號選出適合的真確程度。
3. 有些題目比較相似，故此請小心閱讀每條問題。
4. 請記下開始及完成的時間，並於完成後填於指定欄位上。

	一．問題部份	經常真確	通常真確	有時真確	很少真確	從不真確
1)	我能夠輕易地用單手提起一張木製椅子。					
2)	與一般年齡相近的人比較，我較能輕鬆地完成二十分鐘的中速跑。					
3)	我能夠輕易地以單手拿起一樽 5 公升的食油。					
4)	與身高相若並且同性別的人士比較，我通常比他/ 她們穿較大號的衣服。					
5)	我能輕易以一步跨上三級樓梯（即從地面跨上第三級）。					
6)	我在洗澡時能用左手洗刷腰背所有位置。					
7)	在公共交通工具中，一個座位並不足夠我坐。					
8)	我能夠輕易地以單手拿起一罐 2.5 公升的罐裝食油。					
9)	在超級市場中，我能輕易地伸直手拿取比自己高約 1.5 尺（45 厘米）的物件。					
10)	跪或蹲在地上，然後返回站立姿勢時，我會感到暈眩。					
11)	我坐在沙發上，雙腳平放在地，不用雙手的協助也能輕易地站起來。					
12)	我不能夠輕易地以同一隻手拿著一包八公斤的米超過一分鐘。					
13)	我步行時被後面的人追上。					
14)	我不能夠輕易地以同一隻手拿著一樽 5 公升的食油超過一分鐘。					

		經常真確	通常真確	有時真確	很少真確	從不真確
15)	坐公共交通工具時，我只需一個座位已經十分足夠了。					
16)	於放下重物後（如一包五公斤的米），我便很難再拿起它。					
17)	跪或蹲在地上，然後返回站立姿勢時，我會感到氣喘。					
18)	在沒有任何人或物件的協助下，我能輕易地把背包的兩條肩帶以任何方式放上兩肩。					
19)	兩腿直立時，兩腳掌內側不能緊貼。					
20)	步行時，我會比一般年齡相近的人容易喘氣。					
21)	我能夠輕易地以單手拿起一瓶 6 公升的蒸餾水或礦泉水。					
22)	收拾床舖時，我做出揚起被單的動作沒有感到困難。					
23)	我能夠輕易地以雙手搬起一張單人梳化，並移動兩米距離。					
24)	我不能夠輕易地以同一隻手拿著一排 3 x 900 毫升（共 2.7 公升）的食油超過一分鐘。					
25)	我能輕易地登上公共交通工具（如巴士、小巴、電車等），而不需雙手或他人的協助。					

		經常真確	通常真確	有時真確	很少真確	從不真確
26)	坐在椅上時，挺直腰背、上身與大腿成九十度角的動作，對我來說完全無問題。					
27)	我蹲下繫鞋帶時完全沒有困難。					
28)	我能夠輕易地以單手拿起一包六罐裝(355 毫升 x 6，共 2.13 公升）的汽水或啤酒。					
29)	我平日走路的節奏可以比一般同年紀和同性別的人士快。					
30)	在約十米的距離內急步來回十次後，我能於一分鐘內恢復正常呼吸和心跳。					
31)	我在洗澡時能用右手洗刷腰背所有位置。					
32)	我能夠輕易地以單手拿起一排六包裝（250 毫升 x 6，共 1.5 公升）的紙包飲品。					
33)	與一般年齡相近的人比較，我較能輕鬆地完成二十分鐘的緩步跑。					
34)	我能夠輕易地以單手拿起一袋八公斤的米。					
35)	在沒有任何人或物件的協助下，我能輕易地先把背包的一條肩帶放在其中一肩膀上，再把另一條肩帶放在另一肩膀上。					

		經常真確	通常真確	有時真確	很少真確	從不真確
36)	對我來說，用雙手拿起一個普通大小的圓形西瓜是一件很吃力的事。					
37)	我能夠輕易地單手提著一部手提電腦超過三分鐘而不會感到疲累（以正常的重量計算）。					
38)	我起床時，即使不用手撐著床或扶著任何物件，也能自己坐起來。					
39)	我不能夠輕易地以同一隻手拿著一罐 2.5 公升的罐裝食油超過一分鐘。					
40)	坐在椅上挺直上身時，我的腹部完全不會接觸到大腿。					
41)	我連續步行上三十級樓梯（約一至兩層樓），也不會喘氣。					
42)	我盤膝坐在地上後，只用雙腿的力量也可以站起來。					
43)	我能夠輕易地以單手拿起一袋五公斤的米。					
44)	背部兩肩胛中間發癢時，我能舉起右手抓癢。					
45)	朋友都認爲我身型適中。					
46)	我不能夠輕易地以同一隻手拿著一包五公斤的米超過一分鐘。					
47)	在沒有任何人或物件的協助下，我能夠自行背上背包而不會感到困難。					
		經常真確	通常真確	有時真確	很少真確	從不真確
48)	背部兩肩胛中間發癢時，我能舉起左手抓癢。					
49)	我原地跳時會感到自己很重。					
50)	我的下圍比一般同身高的人士爲大。					
51)	與一般年齡相近的人比較，我快跑 5 至 10 步後不會喘氣。					
52)	跪或蹲在地上，然後返回站立姿勢時，我會心跳加速。					
53)	我不能夠輕易地以同一隻手拿著一包六罐裝（355 毫升 x 6，共 2.13 公升）的汽水或啤酒超過一分鐘。					
54)	我蹲在地上，即使完全不用扶著任何人或物件也能輕易地站起來。					
55)	我的腰圍比一般同身高的人士爲大。					
56)	在巴士、火車或地鐵車廂中，我能完全地舉手緊握車廂天花的扶手。					
57)	以極快速跑約 50 米後（如追巴士），我的心跳和呼吸都能於一分鐘內回復正常。					

		經常真確	通常真確	有時真確	很少真確	從不真確
58)	我的動作較其他年齡相近的人敏捷。					
59)	平日購物後（包括超級市場），我能夠用同一隻手提著約五公斤的貨品步行超過一分鐘。					
60)	我連續步行上三十級樓梯（約一至兩層樓，例如從地下鐵或火車的月台步行上大堂），下肢肌肉會感到疲倦。					
61)	搬動沉重的物件時，我感到很容易拉傷肌肉或關節。					
62)	我能以中等並且保持平均的速度跑畢四百米（圍繞一個標準田徑場一周）。					
63)	我能輕易地用雙手搬起一張有扶手的單座位梳化。					
64)	我能夠輕易地以單手搬動一張木製椅子，並移動最少兩米距離。					
65)	我在平路上急步走一分鐘會喘氣。					
66)	我活動時感到身型影響到我的活動能力和敏捷性。					
67)	後腰中間發癢時，我能用右手從下而上來抓癢。					
68)	我不能夠輕易地以同一隻手拿著一瓶 6 公升的蒸餾水或礦泉水超過一分鐘。					
69)	我保持上身挺直地站立時，低下頭能望到自己的腳趾。					
70)	與一般年齡相近的人比較，我能較輕鬆地快跑 5 至 10 步。					
71)	我能以站立姿勢，彎腰拾起地上不太重的物件而不用屈膝。					
72)	我能夠在巴士車廂中的樓梯輕鬆地上落。					
73)	我坐在椅子上，能夠抱著腳來剪腳甲。					
74)	我盤膝坐地完全沒有困難。					
75)	坐在椅子上，雙腳放在地上，我能輕易地向前彎腰繫鞋帶。					
76)	我能夠輕易地以單手拿起一排 8 罐裝的 355 毫升的汽水或啤酒。					
77)	後腰中間發癢時，我能用左手從下而上來抓癢。					

二・答卷者資料

（請全部填寫）

性別：男 / 女（請圈出適用者）

年齡：_____歲

所需答卷時間：______分鐘

～～～～十分感謝閣下完成問卷！！～～～～

Appendix C

Chinese University of Hong Kong
Department of Sports Science and Physical Education
Self-reported Physical Fitness Questionnaire

Ref. no. : ____________ Date : ________(mm/ dd/ yyyy)

1. The questions focus on issues related to physical fitness which could be encountered in daily lives.
2. Personal information will not be collected.

Target:
Chinese adults aged 24 to 60, residing in Hong Kong.

Instructions:
1. Please consider your physical condition in the previous three months before answering each question. Please select ONE choice only.
2. Each statement describes a hypothetical situation. Please consider the trueness of each statement in describing your physical performance and ability.
3. Please indicate the most suitable choice with a tick (✓) after each statement.
4. Please read every statement very carefully because some statements are very similar.
5. Please indicate on the required space on the last page the time you spend on completing the questionnaire.
6. The questionnaire will approximately take 15 minutes.

	1. Questions					
		Always true	Usually true	Sometimes true	Seldom true	Never true
1)	I can easily RAISE a wooden chair with one arm.					
2)	Compared with people of a similar age, I am more able to easily finish a 20-minute jog at moderate speed.					
3)	I can easily LIFT a bottle of 5-litre vegetable oil with one arm.					
4)	Compared with people of the same sex and similar height, I usually wear larger clothing.					
5)	I can easily go up two steps at a time.					
6)	When taking shower, I can wash anywhere on my back and waist with my left hand.					

		Always true	Usually true	Sometimes true	Seldo m true	Never true
7)	In any kind of public transport, one seat is NOT enough for me.					
8)	I am able to easily LIFT a 2.5-litre canned vegetable oil with one arm.					
9)	In a supermarket, I can easily reach the goods placed 1.5-feet (45 cm) above me.					
10)	I feel dizzy when returning from kneeling or squatting to a normal standing position.					
11)	When sitting on the sofa with legs flattened on the floor, I DO NOT need to use my hand(s) to stand up easily.					
12)	I CANNOT carry a bag of 8-kg rice longer than 1 minute with the same arm.					
13)	People pass me when walking.					
14)	I am NOT able to carry a bottle of 5-kg vegetable oil longer than 1 minute with the same arm.					
		Always true	Usually true	Sometimes true	Seldom true	Never true
15)	In any kind of public transport, one seat is enough for me.					
16)	If I put something heavy down, I find it difficult to pick it up again (e.g. a pack of 5-kg rice).					
17)	I breathe heavily when returning from kneeling or squatting to a normal standing position.					
18)	I can easily place the two straps of a backpack, by any means, onto my shoulders without any assistance.					
19)	Standing with legs together, the inner portions of my feet CANNOT touch each other.					
20)	When walking, I breathe heavily more quickly than other people of the same age.					
21)	I can easily lift a 6-litre water with one arm.					
22)	I find no difficulty in tossing the quilt with both arms.					
23)	I can easily lift a one-seat sofa with both arms and move it a distance of 2 meters.					

		Always true	Usually true	Sometimes true	Seldom true	Never true
24)	I CANNOT able to easily hold a pack of 3-bottle 900ml-vegetable oil (2.7 liters) with one arm for more than 1 minute.					
25)	I can easily get on public transportation (e.g. bus, minibus, and tram) without the assistance of my hands or of anyone.					
26)	I have NO problem in sitting on a chair to an upright position with a right angle between the upper body and the thighs.					
27)	I have NO problem in squatting down to tie up shoelaces.					
28)	I can easily raise a six-can (355ml each) pack of canned drinks (e.g. coke, beer) with one arm.					
29)	Compared with people of the same sex and similar age, I walk faster.					
30)	After walking back and forth ten times over a distance of 10-metre, my breathing and heartbeat can return to normal within 1 minute.					
		Always true	Usually true	Sometimes true	Seldom true	Never true
31)	I can wash all areas of my back with my right hand.					
32)	I can easily raise a six-box (250ml each) pack of drinks (e.g. juice) with one arm.					
33)	Compared with people of a similar age, I am more able to easily finish a 20-minute jogging.					
34)	I can easily lift a bag of 8-kg rice.					
35)	I can easily put one strap of a backpack on a shoulder and the other one on another shoulder without requiring the assistance of anybody or anything.					
36)	Holding a watermelon of normal size is very difficult for me.					
37)	I can easily hold a lap-top computer (normal weight) with one arm for more than 3 minutes without fatigue.					
38)	I can get out of bed without requiring the assistance of my hands or of anything.					

		Always true	Usually true	Sometimes true	Seldom true	Never true
39)	I CANNOT easily hold a 2.5-litre canned oil for more than 1 minute with one arm.					
40)	Sitting in an upright position, my abdomen does not have any contact with my thighs.					
41)	When walking up 30 consecutive steps (1 to 2 floors), I do not breathe heavily.					
42)	I can stand up from a cross-legged sitting position using my legs only.					
43)	I can easily lift a bag of 5-kg rice with one arm.					
44)	When there is an itch between the shoulder blades (scapulas) on of my back, I can scratch with my right hand over the shoulders.					
45)	Friends think I have a medium body size.					
46)	I CANNOT easily hold a bag of 5-kg rice with the same arm for more than 1 minute.					
47)	I can put on the backpack easily without requiring the assistance of anybody or anything.					
		Always true	Usually true	Sometimes true	Seldom true	Never true
48)	When there is an itch between the shoulder blades (scapulas) on of my back, I can scratch with my left hand over the shoulders.					
49)	I feel very heavy when jumping up and down.					
50)	My hip circumference is larger than others of a similar height.					
51)	Compared with other people of similar age, I am less likely to breathe heavily after running 5 to 10 steps fast.					
52)	My heartbeat speeds up when standing up from kneeling or squatting position on the floor.					
53)	I CANNOT easily hold a pack of six 355-ml canned soft drinks/ beer for more than 1 minute with one arm.					

		Always true	Usually true	Sometimes true	Seldom true	Never true
54)	I can stand up easily from a kneeling position without relying on anybody or anything.					
55)	My waist circumference is larger than others of a similar height.					
56)	In a bus or the carriage of a train or MTR, I am able to raise the arm and completely hold the handrails in the ceiling.					
57)	My heartbeat returns to normal within 1 minute after running briskly for approximately 50 meters (e.g. chasing after a bus).					
58)	I am more physically agile than other people of similar age.					
59)	After shopping (including at supermarkets), I can hold goods of 5 kg with the same hand and walk for more than 1 minute.					
		Always true	Usually true	Sometime s true	Seldo m true	Neve r true
60)	Climbing 30 stairs (e.g., 1 to 2 floors from the railway platform to the concourse) makes my legs tired.					
61)	When moving heavy objects, I feel that it is quite likely that I will injure my muscles or joints.					
62)	I can run 400 meters at a medium and steady pace (i.e. same as running around a standard athletic track once).					
63)	I can easily lift a 1-seat sofa (armchair) with two hands.					
64)	I can easily move a wooden chair at least 2 meters using just one arm.					
65)	I breathe heavily when walking fast on level ground for 1 minute.					
66)	I feel that the size of my body has an impact on the ability and agility of my bodily movement.					
67)	When there is an itch at the back of my waist, I can scratch with my right hand.					

		Always true	Usually true	Sometimes true	Seldom true	Never true
68)	I CANNOT easily hold a 6-litre bottle of water with one arm for more than 1 minute.					
69)	Standing upright, I can see my toes by bending my neck forward.					
70)	I can quickly run 5 to 10 steps more easily than others of my age.					
71)	I can bend over, without bending my knees, to pick up objects which are not too heavy on the ground.					
72)	I can move effortlessly in the staircase in a bus.					
73)	I can cut my toenails by sitting on a chair and holding my bent leg.					
74)	I have NO problem sitting cross-legged on the floor.					
75)	Sitting on a chair and leaving both legs on the floor, I can easily bend over to tie up the shoelaces.					
76)	I can easily lift a pack of EIGHT 355-ml canned soft drinks/ beer with the same arm.					
77)	When there is an itch at the back of my waist, I can scratch with my left hand.					

2. Particulars
（Please fill in all blanks.）

Sex：Male / Female（Please circle the suitable one.）

Age：_____ Years

Time used to complete this questionnaire：_____ Minutes

～～Thank you very much for completing this questionnaire！！～～

Appendix D

健康問卷 - HHQ

姓名：________________________年齡: ________ 性別：______聯絡電話: ____________________

I. MEDICAL HISTORY 本人之病歷概況

1. 參考下列之各類健康狀況，如您現在患有任何疾病，請在□內以”✓”表示。如您曾經患有任何疾病，請在□內以”✓”表示，並註明患病之日期。

		現在	曾經（日期）		現在	曾經（日期）
1). Asthma 哮喘		□	□ 由____至____	13). Migraine headache 偏頭痛	□	□ 由____至____
2) .Heart problem 心臟病		□	□ 由____至____	14). Mental disorder 精神/情緒問題	□	□ 由____至____
3). Allergies 過敏病	a）皮膚敏感	□	□ 由____至____	15). Hearing disability 失聰	□	□ 由____至____
		□	□ 由____至____			
	b）鼻敏感	□	□ 由____至____			
	c) 其他敏感					
4). Hepatitis 肝炎		□	□ 由____至____	16). Visual disability 視力衰退	□	□ 由____至____
5). Chest problem 胸肺病		□	□ 由____至____	17). Sinus trouble 鼻竇疾病	□	□ 由____至____
6). Arthritis 關節炎		□	□ 由____至____	18). Anemia 貧血	□	□ 由____至____
7). Diabetes 糖尿病		□	□ 由____至____	19).Thalassaemia 地中海貧血	□	□ 由____至____
8). Back trouble 腰酸背痛		□	□ 由____至____	20). Skin problems 皮膚病	□	□ 由____至____
9). Gynaecological problems 婦科疾病		□	□ 由___至____	21). Tuberculosis 肺結核	□	□ 由____至____
10). Hypertension 高血壓		□	□ 由____至____	22). Blood transfusion 曾接受輸血	□	□ 由____至____
11). High Cholesterol 高膽固醇		□	□ 由____至____	23). AIDS/HIV +ve 愛滋病/愛滋病帶菌者	□	□ 由____至____
12). Kidney disease 腎病		□	□ 由____至____	24). Cancer 癌症 註明________癌	□	□ 由____至____
25). Others 其他 (請註明) ________________						

2. 您曾否接受手術？　　　沒有 □ / 有 □（請列出）

日期	何種手術	入住醫院

3. 您現在有沒有服用藥物（包括中藥、西藥或補充劑）?

□ 沒有 / □有

□ 中藥 □ 西藥 □ 補充劑（請說明藥物資料及其效用）: ____________________

II. 家庭成員

				此項由工作人員填寫
1)	您的以下家庭成員有沒有在 55 歲或以前死於心臟病：	(可選擇多項)	有□ 沒有□	<=55 死於 CHD (有 1 否 0)____
	父親□ 兄弟□ 兒子 □			
2)	您的以下家庭成員有沒有在 65 歲或以前死於心臟病：	(可選擇多項)	有□ 沒有□	<=65 死於 CHD (有 1 否 0)____
	母親□ 姊妹□ 女兒 □			

3). 如上述任何家庭成員曾患有以下疾病，請以"√"表示，並說明患病日期。

		患病日期			患病日期			患病日期
1). Tuberculosis 肺結核	☐	______	2). Kidney disease 腎病	☐	______	3). Hypertension 高血壓	☐	______
4). Epilepsy 癲癇	☐	______	5). Diabetes 糖尿病	☐	______	6). Asthma 哮喘	☐	______
7). Atopic eczema 濕疹	☐	______	8). Nasopharyngeal cancer 鼻咽癌	☐	______	9). Mental disorder 精神/情緒問題	☐	______
10). Heart disease 心臟病	☐							

Appendix E

Physical Activity Readiness Questionnaire

體能活動適應能力問卷與你

(加拿大運動生理學會，2002)

經常進行體能活動不但有益身心，而且樂趣無窮，因此，開始每天多做運動的人愈來愈多。對多數人來說，多做運動是很安全的。不過，有些人在增加運動量前，應先徵詢醫生的意見。如果你計劃增加運動量，請先回答下列7條問題。如果你介乎15-69歲之間，這份體能活動適應能力問卷會告訴你應否在開始前諮詢醫生。<u>如果你超過69歲及沒有經常運動，請徵詢醫生的意見。</u>

普通常識是回答這些問題的最佳指引。請仔細閱讀下列問題，然後誠實回答：
請答「是」或「否」

	是	否	問題
1.	☐	☐	醫生曾否說過你的心臟有問題，以及只可進行醫生建議的體能活動?
2.	☐	☐	你進行體能活動時，是否感到胸口痛?
3.	☐	☐	過去一個月，你曾否在沒有進行體能活動時也感到胸口痛?
4.	☐	☐	你曾否因感到暈眩而失去平衡，或曾否失去知覺?
5.	☐	☐	你的骨骼或關節是否有毛病?且會因改變體能活動而惡化?
6.	☐	☐	醫生現時是否有給你一些有關血壓或心臟藥物(例如去水丸)給你服用?
7.	☐	☐	是否有其他理由令你不應進行體能活動?

如果你的答案「是」:	
一條或以上答「是」	全部答「否」
在開始增加運動量或進行體能評估前，先致電或親身與醫生商談，告訴醫生這份問卷，以及你答「是」的問題。	如果你對體能活動適應能力問卷的全部問題誠實地答「否」，你可合理地相信你可以：
◆ 只要在開始時慢慢進行，然後逐漸增加，你可以進行任何活動；又或者你須受限制，只可進行那些對你安全的活動。告訴醫生你希望參加的活動及聽從他的意見。 ◆ 找出那些對你安全及有幫助的社區活動?	◆ 開始增加運動量……開始時慢慢進行，然後逐漸增加，這是最安全和最容易的方法。 ◆ 參加體能評估，這是一種確定基本體能的好方法，以便你擬定最佳的運動計劃。

延遲增加運動量：
◆ 如果你因傷風或發燒等暫時性疾病而感到不適，請在康復後才增加運動量；或 ◆ 如果你懷孕或可能懷孕，請先徵詢醫生的意見或更改你的體能活動。

加拿大運動生理學會及其代理人毋須爲進行體能活動的人承擔責任。如填妥問卷後有疑問，請先徵詢醫生的意見，然後進行體能活動。

此問卷有效期爲一年

根據加拿大運動生理學會指引，任何人士均歡迎使用及影印此問卷，惟必須整份問卷影印，不能祇選取部份內容。

姓名: ______________	見證人:______________
日期: ______________	(遞交前，請由第三者 與你一起填寫本問卷，然後在此簽署)

Appendix F

Informed Consent

參加《健康及體適能測試》同意書

本人（參加者姓名）＿＿＿＿＿＿＿＿＿＿＿＿＿＿已完全明白「健康及體適能測試」的內容，包括：

一、完成一次健康及體適能測試；及

二、填寫一份體適能自我評估問卷兩次（每次相距至少三天時間）

本人明白參與是次測試純屬自願，亦有權終止。測試負責人保留拒絕或終止不遵守此計劃所訂定之內容和不合適的參加者繼續參與本計劃之權利。本人明白所有收集到的資料或數據，均絕對保密，只供作研究之用。

參加者姓名: ＿＿＿＿＿＿＿＿＿＿＿＿

性別: *男 / 女　　年齡: ＿＿＿＿

聯絡電話號碼: ＿＿＿＿＿＿＿＿

住宅電話號碼: ＿＿＿＿＿＿＿＿

參加者簽署: ＿＿＿＿＿＿＿＿＿＿＿＿　日期: ＿＿＿＿＿＿＿＿＿＿

見證員姓名: ＿＿＿＿＿＿＿＿＿＿＿

見證員簽署: ＿＿＿＿＿＿＿＿＿＿＿　日期: ＿＿＿＿＿＿＿＿＿＿

* 刪去不適用者

Appendix G

Recommended Self-reported Physical Fitness Questionnaire (Women)

	#		Fitness Component	經常真確	通常真確	有時真確	很少真確	從不真確
3)	1	我能夠輕易地以單手拿起一樽 5 公升的食油。	MS					
4)	2	與身高相若並且同性別的人士比較，我通常比他/ 她們穿較大號的衣服。	BC					
6)	3	我在洗澡時能用左手洗刷腰背所有位置。	FLEX					
16)	4	於放下重物後（如一包五公斤的米），我便很難再拿起它。	ME					
28)	5	我能夠輕易地以單手拿起一包六罐裝（355 毫升 x 6，共 2.13 公升）的汽水或啤酒。	MS					
31)	6	我在洗澡時能用右手洗刷腰背所有位置。	FLEX					
34)	7	我能夠輕易地以單手拿起一袋八公斤的米。	MS					
36)	8	對我來說，用雙手拿起一個普通大小的圓形西瓜是一件很吃力的事。	MS					
45)	9	朋友都認爲我身型適中。	BC					
48)	10	背部兩肩胛中間發癢時，我能舉起左手抓癢。	FLEX					
50)	11	我的下圍比一般同身高的人士爲大。	BC					
55)	12	我的腰圍比一般同身高的人士爲大。	BC					
59)	13	平日購物後（包括超級市場），我能夠用同一隻手提著約五公斤的貨品步行超過一分鐘。	ME					
75)	14	坐在椅子上，雙腳放在地上，我能輕易地向前彎腰繫鞋帶。	FLEX					
76)	15	我能夠輕易地以單手拿起一排 8 罐裝的 355 毫升的汽水或啤酒。	MS					

Statistical findings:

Validity (r):	
BC	-.60**
MS	.30**
ME	.20
FLEX	.39**

*$p < .05$; ** $p < .01$

Test-retest reliability (r):	
BC	.90**
MS	.91**
ME	.69**
FLEX	.78**

*$p < .05$; ** $p < .01$

Internal consistency (α):

	Trial 1	Trial 2
BC	.83	.85
MS	.87	.90
ME	.57	.62
FLEX	.74	.81

Appendix H

Recommended Self-reported Physical Fitness Questionnaire (Men)

	#		Fitness Component	經常真確	通常真確	有時真確	很少真確	從不真確
3)	1	我能夠輕易地以單手拿起一樽 5 公升的食油。	MS					
4)	2	與身高相若並且同性別的人士比較，我通常比他/她們穿較大號的衣服。	BC					
8)	3	我能夠輕易地以單手拿起一罐 2.5 公升的罐裝食油。	MS					
18)	4	在沒有任何人或物件的協助下，我能輕易地把背包的兩條肩帶以任何方式放上兩肩。	FLEX					
50)	5	我的下圍比一般同身高的人士爲大。	BC					
55)	6	我的腰圍比一般同身高的人士爲大。	BC					
64)	7	我能夠輕易地以單手搬動一張木製椅子，並移動最少兩米距離。	ME					
66)	8	我活動時感到身型影響到我的活動能力和敏捷性。	BC					
71)	9	我能以站立姿勢，彎腰拾起地上不太重的物件而不用屈膝。	FLEX					

Statistical findings:

Validity (*r*):	
BC	-.57**
MS	.31**
ME	.25
FLEX	.57**

*$p < .05$; ** $p < .01$

Test-retest reliability (*r*):	
BC	.82**
MS	.82**
ME	.67**
FLEX	.78**

*$p < .05$; ** $p < .01$

Internal consistency (α):

	Trial 1	Trial 2
BC	.81	.73
MS	.77	.86
ME	N/A	N/A
FLEX	.32	.28